THE FIELDS GENERATED BY THE VALUES OF THE CHARACTERS OF THE FINITE CLASSICAL GROUPS

John Columba Kelly

Thesis submitted for the degree of Ph. D. at the University of Warwick

Department of Mathematics

September 1975

TABLE OF CONTENTS

Acknowledgements

I wish to thank my research supervisor, Professor George Lusztig, for proposing the problems tackled in this thesis and for frequent help and advice on them. I wish also to thank Professor J.A. Green for a helpful conversation on the content of Section 5.

I was supported by a grant from the Science Research Council while I did this research.

Summary

The absolutely irreducible characters of a finite group G take their values in a cyclotomic extension of the field of rational numbers, and so are defined in some cyclotomic extension L of any field K of characteristic zero. The purpose of this thesis is to determine the smallest such extension L where K is either the rational field or a certain p-adic field, and G is one of the classical groups defined over a finite field with q elements, F_q. In particular it is shown that there is at least one group defined over F_q in each classical "family" whose characters take values in the ring of Witt vectors $W(F_q)$.

Section 5 of the thesis deals with a different problem. A theorem of Frobenius, expressing all the irreducible characters of the symmetric groups as integral linear combinations of the characters of certain permutation representations, is generalised to the Weyl group $W(C_n)$, and some consequences for the representation theory of finite groups with (B,N)-pairs are deduced.

Section 1 Background from Number Theory

1.1 Highest common factors, solution of congruences and Euler's function

If $a_1,a_2,\ldots,a_n$ are (non-zero) positive integers, we write $(a_1,\ldots,a_n)$ for their highest common factor and $[a_1,\ldots,a_n]$ for their least common multiple. The following properties are well known and easily proved.

Proposition 1.1.1

(i) For all positive integers a and b

(a) (a,b) divides ca+db for all integers c and d

(b) Every integral multiple of (a,b) can be written in the form ca−db for some integers c and d, both greater than 0

(c) If a divides e and b divides f, then (a,b) divides (e,f)

(d) If c divides ab, then c/(a,c) divides b.

(ii)(a) (ca,cb) = c(a,b) for all positive integers a, b and c

(b) (ca,b) = (a,b) if (b,c) = 1.

(iii) (a,b) = (b,a) and $[a,b] = [b,a]$

$((a_1,a_2,\ldots,a_{n-1}),a_n) = (a_1,\ldots,a_{n-1},a_n)$

$[[a_1,a_2,\ldots,a_{n-1}],a_n] = [a_1,\ldots,a_{n-1},a_n]$.

(iv) $([a_1,\ldots,a_n],b) = [(a_1,b),(a_2,b),\ldots,(a_n,b)]$

$[(a_1,\ldots,a_n),b] = ([a_1,b],[a_2,b],\ldots,[a_n,b])$.

(v) $(a,b).[a,b] = ab$.

Proof Omitted.

If a, b and c are integers with $c>0$, we write the "congruence"

$$a \equiv b \bmod(c)$$

to mean that c divides a−b. The following result is a form of the Chinese Remainder Theorem. It is given here in the form most suitable for later computations but will be restated in Lemma 1.1.4(ii) in a

more concise algebraic form.

Theorem 1.1.2

There is a solution for x of the n simultaneous congruences

$$x \equiv a_i \bmod(b_i) \text{ for } i = 1,2,\ldots,n$$

if and only if $a_i \equiv a_j \bmod((b_i,b_j))$ for each i and j. If this holds, the n simultaneous congruences are equivalent to the single congruence

$$x \equiv a \bmod([b_1, \ldots, b_n])$$

where a is any solution of the original congruences.

Proof The necessity of the conditions $a_i \equiv a_j \bmod((b_i,b_j))$ is obvious. We prove sufficiency by induction on n. The result is trivial for n=1.

Suppose that n = 2. We wish to solve

$$x \equiv a_1 \bmod(b_1) \text{ and } x \equiv a_2 \bmod(b_2)$$

where $a_1 \equiv a_2 \bmod((b_1,b_2))$. By Proposition 1.1.1(i)(b) $a_1-a_2=cb_1-db_2$ with c and d integers. Then $x = a = a_1+cb_1 = a_2+db_2$ is a solution of the simultaneous congruences. If $x = a'$ is another solution, then $a-a' \equiv 0 \bmod(b_i)$ for i = 1,2 and so $[b_1,b_2]$ divides a-a'. Thus the simultaneous congruences are equivalent to the single one:

$$x \equiv a \bmod([b_1,b_2]) \text{ as required.}$$

Suppose next that $n>2$ and that the theorem holds for n-1 congruences. Thus the first n-1 congruences are equivalent to

$$x \equiv a' \bmod([b_1,\ldots,b_{n-1}]) \text{ where } a' \equiv a_i \bmod(b_i) \text{ for } 1 \leqslant i \leqslant n-1.$$

Hence $a' \equiv a_i \bmod((b_i,b_n))$ for $1 \leqslant i \leqslant n-1$

$\equiv a_n \bmod((b_i,b_n))$ by hypothesis.

Hence $a' \equiv a_n \bmod([(b_1,b_n),\ldots,(b_{n-1},b_n)])$,

i.e. $a' \equiv a_n \bmod(([b_1,\ldots,b_{n-1}],b_n))$ by Proposition 1.1.1(iv).

The theorem follows by applying the case n = 2 to the two congruences: $x \equiv a' \bmod([b_1,\ldots,b_{n-1}])$ and $x \equiv a_n \bmod(b_n)$, and using Proposition 1.1.1(iii).

Note The Chinese Remainder Theorem is usually given as Theorem 1.1.2

in the special case where $(b_i,b_j) = 1$ for $i \neq j$. The form given above is a special case of Theorem 17, Chapter V of Zariski and Samuel's "Commutative Algebra" where it is stated in terms of ideals and proved for a class of rings which includes Dedekind domains. The theorem is false in general for unique factorisation domains.

In applications of this result we shall often need highest common factors of the following forms.

Proposition 1.1.3

Let q, i and j be positive integers with $q>1$. Then

(i) $(q^i-1,q^j) = (q^i+1,q^j) = 1$

(ii) $(q^i-1,q^j-1) = q^{(i,j)}-1$

(iii) $(q^i+1,q^j+1) = \left\{\begin{matrix} q^{(i,j)}+1 & \text{if } v_2(i) = v_2(j) \\ (2,q+1) & \text{if } v_2(i) \neq v_2(j) \end{matrix}\right\}$ where $v_2(k)$ is the highest power of 2 dividing k.

(iv) $(q^i-1,q^j+1) = \left\{\begin{matrix} q^{(i,j)}+1 & \text{if } v_2(i)>v_2(j) \\ (2,q+1) & \text{otherwise} \end{matrix}\right\}$

Proof (i) Apply Proposition 1.1.1(i) to the identity

$$(-(1+q^i+q^{2i}+\ldots+q^{(j-1)i}))(q^i-1) + q^{(i-1)j}q^j = 1.$$

(ii) Choose positive integers c and d with $(i,j) = ci-dj$. Then (q^i-1,q^j-1) divides $(q^{ci}-1,q^{dj}-1)$ by Proposition 1.1.1(i)(c), and so divides $(q^{ci}-1)-(q^{dj}-1) = q^{dj}(q^{(i,j)}-1)$. Since $(q,q^i-1) = 1$, the required highest common factor must divide $q^{(i,j)}-1$ which clearly is a factor of both q^i-1 and q^j-1.

(iii) If $v_2(i) = v_2(j)$, then by Proposition 1.1.1(i)(b), $(i,j)=ci-dj$ for some positive integers i and j, exactly one of which is odd, since the same highest power of 2 divides i, j and (i,j). Suppose, without loss of generality, that d is odd. Then (q^i+1,q^j+1) divides $q^{ci}-1+q^{dj}+1 = q^{dj}(q^{(i,j)}+1)$. The proof continues as for (ii).

If $v_2(i) \neq v_2(j)$ we may find c and d with $ci = dj$ and only one of c,d even. If, without loss of generality, c is even then

(q^i+1, q^j+1) divides $-(q^{ci}-1)+(q^{dj}+1) = 2$. The result follows by separate consideration of the cases where q is even and odd.

(iv) If $v_2(i) > v_2(j)$, then there are positive integers c and d with $(i,j) = ci-dj$ and d necessarily odd. Then (q^i-1, q^j+1) divides $q^{ci}-1+q^{dj}+1$ and the proof continues as for (ii).

If $v_2(i) \leqslant v_2(j)$ there are positive integers c,d with d odd and $ci = dj$. Then (q^i-1, q^j+1) divides $-(q^{ci}-1)+(q^{dj}+1) = 2$. The proof continues as for (iii) above.

Euler's function $\phi(m)$ is defined for positive integers m as follows: $\phi(1) = 1$ by convention

$\phi(m)$ = the number of residues modulo m which are prime to m for $m > 1$.

Alternatively (and more algebraically) it may be defined by

$\phi(m) = |(Z/mZ)^*|$, where $(Z/mZ)^*$ is the group of units of the finite commutative ring Z/mZ.

Lemma 1.1.4

(i) If m divides n, the natural map $f_{nm}: Z/nZ \longrightarrow Z/mZ$ given by $f_{nm}(a+nZ) = a+mZ$ is a surjective ring homomorphism which restricts to an epimorphism of the groups of units.

(ii) The natural map $f: Z/[b_1,\ldots,b_n]Z \longrightarrow (Z/b_1Z) \times \cdots \times (Z/b_nZ)$ given by $f(a+[b_1,\ldots,b_n]Z) = (a+b_1Z,\ldots,a+b_nZ)$ is an injective ring homomorphism with image the set of $(a_1+b_1Z,\ldots,a_n+b_nZ)$ such that

$$a_i \equiv a_j \bmod((b_i,b_j)) \text{ for all i and j.}$$

Proof (i) f_{nm} is certainly a surjective ring homomorphism. To prove the assertion about the groups of units we must show that given a with $(a,m) = 1$, there exists d with $(a+dm,n) = 1$.

Let $p_1,\ldots,p_k$ be all the distinct prime factors of n. No p_i divides both a and a+m since $(a,m) = 1$, and so for each i there is a $d_i = 0$ or 1 such that p_i does not divide $a+d_im$. By Theorem 1.1.2

there is a d with $d \equiv d_i \bmod(p_i)$ for each i, and so

$a+dm \equiv a+d_i m \not\equiv 0 \bmod(p_i)$ for each i, showing that $(a+dm,n)=1$.

(ii) f is certainly a ring homomorphism. The rest is a translation of Theorem 1.1.2.

Euler's function satisfies a multiplicative property which will be needed in an unusual form. (See Corollary 1.1.6).

Theorem 1.1.5

For each pair of positive integers m and n there is a short exact sequence of abelian groups and homomorphisms:

$$1 \longrightarrow (Z/[m,n]Z)^* \xrightarrow{f} (Z/mZ)^* \times (Z/nZ)^* \xrightarrow{g} (Z/(m,n)Z)^* \longrightarrow 1$$

where $f(a+[m,n]Z) = (a+mZ, a+nZ)$ and $g(a+mZ, b+nZ) = (a+(m,n)Z)(b+(m,n)Z)^{-1}$.

Proof f is an injective and g a surjective group homomorphism by Lemma 1.1.4. Ker(g) is the set of $(a+mZ, b+nZ)$ with $a \equiv b \bmod((m,n))$, which is the same as Im(f) by Lemma 1.1.4(ii). (Since the map of Lemma 1.1.4(ii) is injective, the inverse image of the group of units in the image ring must be the group of units of $Z/[m,n]Z$).

Note The sequence splits.

Corollary 1.1.6

For all positive integers m and n

$$\phi((m,n))\phi([m,n]) = \phi(m)\phi(n).$$

Proof Inspect the orders of the finite groups in the exact sequence.

$\phi(m)$ may be calculated by factorising m into a product of prime powers, using this result, and noting that

$\phi(p^a) = p^a - p^{a-1}$ for p prime and $a \geq 1$.

Note The short exact sequence of Theorem 1.1.5 may be reinterpreted as a sequence of Galois groups. In the notation explained in 1.2

(below) it is the same as the sequence

$$1 \to \mathrm{Gal}(Q(\sqrt[[m,n]]{1})/Q) \xrightarrow{f} \mathrm{Gal}(Q(\sqrt[m]{1})/Q) \times \mathrm{Gal}(Q(\sqrt[n]{1})/Q) \xrightarrow{g} \mathrm{Gal}(Q(\sqrt[(m,n)]{1})/Q) \to 1$$

where $f(s)$ = (restriction of s, restriction of s) and

$g(s,t)$ = (restriction of s)(restriction of t)$^{-1}$.

1.2 Galois theory and cyclotomic fields

If L/K is a field extension of finite degree write

Gal(L/K) = the group of automorphisms of L which fix K pointwise, (the Galois group of L over K).

Let K be a field of characteristic zero, and m be any positive integer. Write $L = K(\sqrt[m]{1})$ for the splitting field over K of the polynomial x^m-1. If ζ is any primitive m^{th} root of unity in L, then the zeros of x^m-1 are precisely $1, \zeta, \zeta^2, \ldots, \zeta^{m-1}$ and $L = K[\zeta]$. If $g \in Gal(L/K)$ then $g\zeta$ must be another primitive m^{th} root of unity, and so $g\zeta = \zeta^k$ for some integer k with (k,m) = 1. It is easily checked that the map $g \longmapsto k$ gives a canonical injection (i.e. independent of the choices of ζ and k) of Gal(L/K) into $(Z/mZ)^*$, and that this injection is a group homomorphism. In particular $[K(\sqrt[m]{1}):K]$ divides $\phi(m)$ and the Galois group of this extension is abelian. Hence all fields lying between K and $K(\sqrt[m]{1})$ are normal (separable) extensions of K.

Theorem 1.2.1

(i) $Q(\sqrt[m]{1})$ is a normal extension of Q of degree $\phi(m)$. Its Galois group is naturally isomorphic to $(Z/mZ)^*$.

(ii) A basis for the Z-module of algebraic integers in $Q(\sqrt[m]{1})$ is $1, \zeta, \zeta^2, \ldots, \zeta^{\phi(m)-1}$ where ζ is a primitive m^{th} root of unity.

Proof See (for example) B.J. Birch: "Cyclotomic fields and Kummer extensions", Chapter III in Cassels & Frohlich (ed.):"Algebraic Number Theory", Academic Press (1967).

Lemma 1.2.2

If L is a normal separable extension of M and K an arbitrary extension, all contained inside some larger field, then $Gal(LK/K) \cong Gal(L/L \cap K)$.

Hence $[LK:K] = [L:L\cap K]$.

Proof This is a special case of a standard result in Galois theory. See (for example) Lang: "Algebra", Chapter VIII, Section 1, Theorem 4.

Proposition 1.2.3

In any algebraic closure of Q, (i) $Q(\sqrt[m]{1})Q(\sqrt[n]{1}) = Q(\sqrt[[m,n]]{1})$

(ii) $Q(\sqrt[m]{1})\cap Q(\sqrt[n]{1}) = Q(\sqrt[(m,n)]{1})$

Proof (i) If ζ is a primitive m^{th} root of unity and η a primitive n^{th} root of unity, then $\zeta\eta$ is a primitive $[m,n]^{th}$ root of unity.

(ii) Applying Lemma 1.2.2 with $L = Q(\sqrt[m]{1})$, $K = Q(\sqrt[n]{1})$ gives

$$\begin{aligned}[Q(\sqrt[m]{1}):L\cap K] &= [Q(\sqrt[[m,n]]{1}):Q(\sqrt[n]{1})] \text{ using (i)}\\ &= \phi([m,n])/\phi(n) \text{ by Theorem 1.2.1(i)}\\ &= \phi(m)/\phi((m,n)) \text{ by Corollary 1.1.6}\\ &= [Q(\sqrt[m]{1}):Q(\sqrt[(m,n)]{1})] \text{ by Theorem 1.2.1(i).}\end{aligned}$$

The result follows since clearly $L\cap K \supseteq Q(\sqrt[(m,n)]{1})$.

In order to manipulate cyclotomic fields later we will need some technical lemmas on Galois groups. If H is a subgroup of Gal(L/K) and M is a subfield of L we write M^H for the subfield of points of M fixed by each element of H. If $g\in Gal(L/K)$ we write M^g for M^H where H is the cyclic subgroup generated by g.

Proposition 1.2.4

Let L and K be finite extensions of M both contained in some larger field. If H is a subgroup of Gal(LK/M) then

(i) $L\cap K^H = (L\cap K)^H$

(ii) If L is a normal separable extension of M and g restricts to the identity map on K for all $g\in H$, (i.e. $K^H = K$), then

$$(LK)^H = L^H K \quad \text{and} \quad [LK:L^H K] = [L:L^H].$$

Proof (i) Obvious.

(ii) By Theorem 2, page 194, Chapter VIII of Lang's "Algebra",

$[LK:(LK)^H]$ = the order of H.

Since L is a normal separable extension of M and each $g \in H$ acts trivially on K, H must act faithfully as a group of automorphisms on L. Hence by the same theorem, the order of H is equal to $[L:L^H]$, which is greater than or equal to $[LK:L^HK]$ by general field theory. However $L^HK \subseteq (LK)^H$ since H acts trivially on K. The result follows immediately.

Proposition 12.5

If L and K are normal separable extensions of M of finite degree and are both contained in some larger field, and $g \in Gal(LK/M)$, then

$$[(LK)^g:L^gK^g] = ([L:L^g],[K:K^g])/[L \cap K:(L \cap K)^g].$$

Proof Clearly $(LK)^g \supseteq L^gK^g$. Now

$$[(LK)^g:L^gK^g] = [LK:LK^g][LK^g:L^gK^g]/[LK:(LK)^g].$$

However $[LK^g:L^gK^g] = [L:L^g]$ by Proposition 1.2.4(ii) with H the cyclic group generated by g, and $[LK:(LK)^g]$ = the order of g acting on LK by Theorem 2, Chapter VIII of Lang's "Algebra", and this is the least common multiple of the orders of g acting on L and on K, that is the least common multiple of $[L:L^g]$ and $[K:K^g]$ by the same theorem. Hence, using Proposition 1.1.1(v),

$$[LK:(LK)^g] = [L:L^g][K:K^g]/([L:L^g],[K:K^g]) \text{ and so}$$

$$[(LK)^g:L^gK^g] = ([L:L^g],[K:K^g]).[LK:LK^g]/[K:K^g].$$

Now $[K:K^g]/[LK:LK^g] = [K:K^g][LK^g:K^g]/[LK:K^g] = [LK^g:K^g]/[LK:K]$

$$= [L:L \cap K^g]/[L:L \cap K] \text{ using Lemma 1.2.2 twice}$$

$$= [L \cap K:L \cap K^g] = [L \cap K:(L \cap K)^g] \text{ using}$$

Proposition 1.2.4(i). This completes the proof.

Corollary 12.6

With the hypotheses of Proposition 1.2.5, $(LK)^g = L^gK^g$ if and only if the order of g on $L \cap K$) (order of g on L, order of g on K).

Proof Substituting this inequality in the statement of Proposition 1.2.5, we find that the field index is an integer ≤ 1. In particular we see that equality must hold if the inequality does.

Proof Substituting this inequality in the statement of Proposition 1.2.5, we find that the field index is an integer $\leqslant 1$. In particular we see that equality must hold if the inequality does.

1.3 Witt vectors and p-adic fields

A (non-archimedean) valuation on a (commutative) integral domain R is a map $|\ |: R \longrightarrow$ the real numbers, written $x \mapsto |x|$, such that (i) $|x| \geqslant 0$ and $|x| = 0$ if and only if $x = 0$

(ii) $|xy| = |x||y|$ for all $x,y \in R$

(iii) $|x+y| \leqslant \max(|x|,|y|)$ for all $x,y \in R$.

R becomes a metric space if we define $d(a,b) = |a-b|$. R is said to be complete (with respect to $|\ |$) if this makes it a complete metric space. Two valuations, $|\ |_1$ and $|\ |_2$, on R are said to be equivalent if $|x|_2 = |x|_1^t$ for some fixed t and all $x \in R$. (In fact it may be shown that this happens if and only if the two valuations define the same topology on R). A valuation $|\ |$ is called discrete if the set of all $|x|$ for $0 \neq x \in R$ is a discrete subset (and hence cyclic subgroup) of the positive real numbers.

A discrete valuation ring is an integral domain R with a discrete non-archimedean valuation $|\ |$ such that $|x| \leqslant 1$ for all $x \in R$, and $|x| = 1$ if and only if x is a unit in R. In this case R is a local ring with unique maximal ideal $\mathfrak{m}$, the set of x with $|x| < 1$, called the valuation ideal of R. R is also a principal ideal domain. $R/\mathfrak{m}$ is called the residue class field of R. The equivalence class of $|\ |$ is completely determined by $\mathfrak{m}$. In what follows we do not distinguish between equivalent valuations.

Let k be a field of characteristic $p \neq 0$. A ring of Witt vectors for k is a pair (R,i) such that

W1. R is a complete discrete valuation ring of characteristic 0

W2. The maximum ideal of R is pR

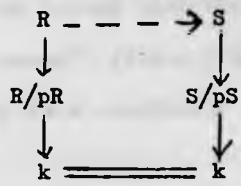
W3. i is a field isomorphism $i: R/pR \xrightarrow{\sim} k$

Theorem 1.3.1

Let k be a perfect field of characteristic $p \neq 0$.

(i) There exists a ring of Witt vectors (R,i) for k.

(ii) If (R,i) and (S,j) both satisfy conditions W1, W2 and W3, then there is a unique isomorphism between R and S making the diagram on the right commute. We write W(k) for the (essentially) unique ring R of Witt vectors for k.



(iii) The correspondence $k \longmapsto W(k)$ is a functor from the category of perfect fields of non-zero characteristic and monomorphisms to the category of complete discrete valuation rings and monomorphisms.

Proof See Chapter 6 of M.J. Greenberg: "Lectures on forms in many variables", Benjamin (1969) as follows:

(i) See the construction on pages 81-91.

(ii) See Theorem 6.3, page 80.

(iii) Clear from the construction used in proving (i)

The geheral construction of Witt vectors used in part (i) of the proof of this theorem does not make routine arithmetic in W(k) very easy. However our main interest will be in the case $k = F_q$, the finite field with $q = p^n$ elements where p is prime. For this case we may use properties of p-adic numbers to give an alternative construction of $W(F_q)$.

Let p be a prime number and Q the field of rational numbers. We may define a non-archimedean valuation $|\ |_p$ on Q by setting $|0|_p = 0$ and $|a/b|_p = p^{m-n}$ where p^n (p^m respectively) is the highest power of p dividing a (b respectively). Let Q_p be the metric space completion of Q with respect to this valuation. The following results are standard.

Theorem 1.3.2

(i) Q_p has a unique field structure extending that of Q such that the field operations $(x,y) \longmapsto xy$, $(x,y) \longmapsto x+y$ and $x \longmapsto x^{-1}$ (for $x \neq 0$) are continuous functions. $|\ |_p$ extends uniquely to a continuous non-archimedean valuation on Q_p.

(ii) Z_p = the set of $x \in Q_p$ with $|x|_p \leqslant 1$ is a complete discrete valuation ring with field of fractions Q_p.

(iii) The valuation ideal of Z_p is pZ_p and the residue class field is F_p with residue class representatives $0,1,\ldots,p-1$ as in Z/pZ.

(iv) Every element a of Z_p may be written uniquely in the form $a = a_0 + a_1 p + \ldots + a_i p^i + \ldots$ with $0 \leqslant a_i \leqslant p-1$, an integer, and the obvious rules of multiplication and addition. (Subtraction may be reduced to these by noting that $-1 = (p-1) + (p-1)p + (p-1)p^2 + \ldots.$). If $a \neq 0$, then $|a|_p = p^{-i}$ where i is the least integer with $a_i \neq 0$. Every element of Q_p may be written as $p^{-n}a$ for some $a \in Z_p$, $n \in Z$.

Proof Omitted.

In particular (ii) and (iii) show that $W(F_p) = Z_p$. We next construct $W(F_q)$ by examining cyclotomic extensions of Q_p.

If F is a field complete with respect to a discrete non-archimedean valuation $|\ |$, the subring of all $x \in F$ with $|x| \leqslant 1$ is called the valuation ring (or ring of integers) of F, and is a complete discrete valuation ring with F as field of fractions. Its residue class field is called the residue class field of F.

Theorem 1.3.3

If F is a field complete with respect to a discrete non-archimedean valuation, $|\ |_F$, and E is an extension of F of degree n, then

(i) $|x|_E = (|N_{E/F}x|_F)^{1/n}$ is the unique non-archimedean valuation on E extending the one on F, where $N_{E/F}: E \to F$ is the norm map.

(ii) E is complete with respect to this valuation which is discrete.

(iii) The valuation ring of E is the integral closure of that of F in E.

(iv) The inclusion map of valuation rings induces a monomorphism

residue class field of F $\longrightarrow$ residue class field of E.

Proof See, for example, O.T. O'Meara: "Introduction to Quadratic Forms", Section 14, Springer-Verlag (1963) for the existence and uniqueness of the extension, section 16 for its discreteness, the proof of existence and uniqueness for part (iii), and then (iv) follows from the others by elementary algebra.

A finite extension E of Q_p is said to be unramified if

valuation ideal of E = p.(valuation ring of E)

with respect to the unique extension of the p-adic valuation, $|\ |_p$, to E. Otherwise the extension is said to be ramified. Subextensions of unramified extensions are unramified.

Theorem 1.3.4

For each positive integer f

(i) there is a unique unramified extension E of Q_p of degree f,

(ii) it is the splitting field of the polynomial $x^{q-1}-1$ where $q = p^f$,

(iii) $Gal(E/Q_p)$ is cyclic of order f generated by the unique automorphism σ such that $|\sigma a - a^p|_E < 1$ for all a with $|a|_E \leq 1$.

(iv) The residue class field of E is F_q with residue class representatives $0, \zeta, \zeta^2, \ldots, \zeta^{q-2}, \zeta^{q-1} = 1$, where ζ is a primitive $(q-1)^{th}$ root of unity in E. The automorphism σ of (iii) acts by

(v) The only roots of unity in E of degree prime to p are the $(q-1)^{th}$ ones.

Proof See section 32 of the book by O'Meara. The condition "of degree prime to p" in (v) can be omitted by the next theorem.

In the notation of this theorem we write $E = Q_q$ and write Z_q for the valuation ring of Q_q. (Both of these notations are non-standard, but they seem quite appropriate in this context). By parts (i) and (iv) of the theorem $Z_q = W(F_q)$. It does not seem possible to give such a simple description of the field operations in Q_q as in Theorem 1.3.2(iv), but we do not need this.

Next we examine cyclotomic extensions of Q_q.

Theorem 1.3.5

(i) There is a unique unramified extension of Q_q of degree f, the field Q_r, the splitting field of the polynomial $x^{r-1}-1$, where $r = q^f$.

(ii) If $(m,q) = 1$, $Q_q(\sqrt[m]{1})$ is an unramified extension of Q_q of degree f = the least positive integer such that m divides q^f-1. (In fact f divides $\phi(m)$).

(iii) If $n = p^a m$ with m and f as in (ii), and q a power of the prime p, then

(a) $[Q_q(\sqrt[n]{1}):Q_q] = f.\phi(p^a)$

(b) If ζ is a primitive m^{th} root of unity, and η a primitive p^ath root of unity, then $Gal(Q_q(\sqrt[n]{1})/Q_q)$ is the internal direct sum of the cyclic subgroups with generators as follows:

for p odd, σ_{ram} and σ_{Fr} such that

$$\left.\begin{aligned}\sigma_{Fr}(\zeta^i\eta^j) &= \zeta^{iq}\eta^j\\ \sigma_{Fram}(\zeta^i\eta^j) &= \zeta^i\eta^{kj}\end{aligned}\right\}$$ for all i and j, where k is a generator of $(Z/p^aZ)^*$ which is cyclic by elementary number theory,

for $p = 2$, $q \geqslant 8$, σ_1, σ_2 and σ_{Fr} where σ_{Fr} is as above, $\sigma_1(\zeta^i\eta^j) = \zeta^i\eta^{-j}$ and $\sigma_2(\zeta^i\eta^j) = \zeta^i\eta^{5j}$ for all i,j.

Note We write G_{ram} for the subgroup generated by σ_{ram} or by σ_1 and σ_2 as appropriate, and G_{Fr} for the cyclic subgroup generated by σ_{Fr}. The Galois group is always the direct sum of these two subgroups, but

for $q = 2$ G_{ram} has only one element, and for $q = 4$ it has order 2. "Fr" and "ram" are abbreviations of "Frobenius" and "ramified".

Proof (i) is an immediate consequence of Theorem 1.3.4(i) and (ii).

(ii) Reducing modulo m, q gives an element of the group $(Z/mZ)^*$, and so there is a least integer f (necessarily dividing $\phi(m)$) such that $q^f \equiv 1 \bmod(m)$. Hence m divides $q^f - 1 = r - 1$ in the notation of (i), and so $Q_q(\sqrt[m]{1}) \subseteq Q_r$ and must be unramified of degree dividing f. The result follows by (i) and Theorem 1.3.4(v).

(iii)(a) By (ii) it is enough to show that a primitive p^ath root of unity η has degree $\phi(p^a)$ over Q_r for any $r = p^b$. Let $\lambda = \eta - 1$. Now $\eta^{p^a} = 1 \neq \eta^{p^{a-1}}$, and so λ is a root of the polynomial

$$h(x) = ((x+1)^{p^a} - 1)/((x+1)^{p^{a-1}} - 1).$$

By inspection $\deg(h) = p^a - p^{a-1} = \phi(p^a)$, and h has leading coefficient 1, constant term p, and all other coefficients divisible by p. So h(x) is an "Eisenstein" polynomial, and the result follows by 32:15 of O'Meara's book.

(iii)(b) By the proof of (iii)(a)

$$Q_q(\sqrt[n]{1}) = Q_q(\sqrt[m]{1})Q_q(\sqrt[p^a]{1}),\ Q_q(\sqrt[m]{1}) \cap Q_q(\sqrt[p^a]{1}) = Q_q.$$

The result is now a straightforward application of Theorem 5 of section 1, Chapter VIII in Lang's "Algebra". (For $p = 2$, we note that $(Z/2^aZ)^*$ is generated by -1 and 5).

Section 2 The Field Generated by the Characters of a Finite Group

Throughout this section G is a finite group, and m is a positive integer with $g^m = 1$ for all $g \in G$. (For example, m could be the order of G or any other multiple of the exponent of G). All representations of G considered are over fields of characteristic zero.

2.1 K(G), $\Gamma(G)$ and $\Gamma(G,K)$

Let L be a field of characteristic zero, V a vector space of dimension n over L and $R: G \longrightarrow GL(V)$ a representation of G in the automorphism group of V, with character χ. Then for each $g \in G$

$$R(g)^m = R(g^m) = R(1) = I,$$

and so the eigenvalues $w_1, \ldots, w_n$ of $R(g)$ are m^{th} roots of unity in some algebraic extension of L. Hence

$$\chi(g) = w_1 + \ldots + w_n \in Q(\sqrt[m]{1})$$

where Q is taken as the prime field of L.

Definitions (i) If K is a field of characteristic zero, define

K(G) = the subfield of $K(\sqrt[m]{1})$ generated over K by the numbers $\chi(g)$ for all $g \in G$ and all absolutely irreducible characters χ of G.

(ii) $\Gamma(G)$ = the set of k in $(Z/mZ)^*$ such that g is conjugate to g^k for all $g \in G$.

(iii) $\Gamma(G,K) = \Gamma(G) \cap Gal(K(\sqrt[m]{1})/K)$ where the Galois group is naturally identified with a subgroup of $(Z/mZ)^*$ as at the beginning of 1.2.

Notes (i) The absolutely irreducible characters of G are defined in $Q(\sqrt[m]{1})$ and do not depend on the choice of the field K.

(ii) g^k is well-defined for $g \in G$ and $k \in (Z/mZ)^*$ since $g^m = 1$.

(iii) It is easily seen that $\Gamma(G)$ is a subgroup of $(Z/mZ)^*$, and so that $\Gamma(G,K)$ is a subgroup of $Gal(K(\sqrt[m]{1})/K)$.

These definitions are linked by:

Theorem 2.1.1

$$K(G) = K(\sqrt[m]{1})^{\Gamma(G,K)} \quad \text{(field of fixed points)}$$

Proof Certainly $K(G) = K(\sqrt[m]{1})^{H}$ for some unique subgroup H of $\mathrm{Gal}(K(\sqrt[m]{1})/K)$. Let R be a representation of G with character χ and $g \in G$ such that $R(g)$ has eigenvalues $w_1, \ldots, w_n$. Then for $k \in (Z/mZ)^*$, $R(g^k) = R(g)^k$ has eigenvalues $w_1^k, \ldots, w_n^k$ and so $\chi(g) = w_1+\ldots+w_n$ and $\chi(g^k) = w_1^k+\ldots+w_n^k = \chi(g)$ acted upon by the element k of $\mathrm{Gal}(Q(\sqrt[m]{1})/Q) = (Z/mZ)^*$.

In particular for $k \in \Gamma(G)$, $\chi(g) = \chi(g^k)$ and so is invariant under the action of k. Hence $\Gamma(G,K) \subseteq H$.

Conversely let k be in $\mathrm{Gal}(K(\sqrt[m]{1})/K)$ but not in $\Gamma(G,K)$. Then there is some $g \in G$ which is not conjugate to g^k, and hence some absolutely irreducible character χ of G such that $\chi(g) \neq \chi(g^k)$. Hence k is not in H, and so $H \subseteq \Gamma(G)$.

Corollary 2.1.2

$K(G) = K$ if and only if $\Gamma(G,K) = \mathrm{Gal}(K(\sqrt[m]{1})/K)$,

if and only if $\Gamma(G) \supseteq \mathrm{Gal}(K(\sqrt[m]{1})/K)$.

Proof obvious by Galois theory and the definition of $\Gamma(G,K)$.

Corollary 2.1.3

$$[K(G):K] = [\mathrm{Gal}(K(\sqrt[m]{1})/K):\Gamma(G,K)] .$$

Proof Obvious by Galois theory.

Note The treatment above follows that of J.P. Serre: "Representations Lineaires des Groupes Finis", Partie II, Section 12.3 (Hermann 1967).

2.2 Cyclotomic extensions of the rational and p-adic fields

If n divides m, let $f_{mn}:(Z/mZ)^* \longrightarrow (Z/nZ)^*$ be defined as in Lemma 1.1.4(i). In particular

$\mathrm{Ker}(f_{mn})$ = the set of $k \in (Z/mZ)^*$ with $k \equiv 1 \bmod(n)$.

Theorem 2.2.1

(i) $Q(G) = Q$ if and only if $\Gamma(G) = (Z/mZ)^*$.

(ii) $Q(G) \subseteq Q(\sqrt[n]{1})$ if and only if $\Gamma(G) \supseteq \mathrm{Ker}(f_{m,(m,n)})$ and in that case $Q(G) = Q(\sqrt[(m,n)]{1})^H$ where $H = f_{m,(m,n)}(\Gamma(G))$.

Proof (i) is a special case of (ii).

(ii) Certainly $Q(G) \subseteq Q(\sqrt[n]{1})$ if and only if $Q(G)$ is contained in $Q(\sqrt[m]{1}) \cap Q(\sqrt[n]{1}) = Q(\sqrt[(m,n)]{1})$ by Proposition 1.2.3, and so it is enough to consider the case "n divides m" and we must prove that $\mathrm{Gal}(Q(\sqrt[m]{1})/Q(\sqrt[n]{1})) = \mathrm{Ker}(f_{mn})$.

Let w be a primitive n^{th} root of unity, and $k \in (Z/mZ)^*$. Then $w^k = w$ if and only if $k \equiv 1 \bmod(n)$, i.e. $k \in \mathrm{Ker}(f_{mn})$.

If $m = p^a n$ with $(p,n) = 1$, and q is a power of the prime p, then by Theorem 1.3.5 $\mathrm{Gal}(Q_q(\sqrt[m]{1})/Q_q) = G_{ram} \oplus G_{Fr}$, where G_{ram} = the set of $k \in (Z/mZ)^*$ with $k \equiv 1 \bmod(n)$, and G_{Fr} is generated by the unique $k \in (Z/mZ)^*$ with $k \equiv 1 \bmod(p^a)$ and $k \equiv q \bmod(n)$.

With this notation the results of 2.1 give

Theorem 2.2.2

(i) $Q_q(G)$ is an unramified extension of Q_q if and only if

$$\Gamma(G) \supseteq \mathrm{Ker}(f_{mn}) = G_{ram}.$$

(ii) $Q_q(G) = Q_q$ if and only if $\Gamma(G) \supseteq G_{ram}$ and G_{Fr}.

Proof A translation of the results of 2.1 using Theorem 1.3.5.

Notes (i) The symmetric groups satisfy Theorem 2.2.1(i). The theorem may also be verified for the other families of Weyl groups.

(ii) Theorem 2.2.2(i) is satisfied (vacuously) if p does not divide m.

(iii) It might be hoped that when the prime power q is "naturally" associated with the group G, then the field $Q_q(G)$ might have a very simple description. Later results will confirm this in many cases where G is a Chevalley group defined over a finite field F_q.

2.3 Relationships with subgroups and homomorphisms

Proposition 2.3.1

Let G and H be finite groups, and K a field of characteristic zero.

(i) $K(G \times H) = K(G)K(H)$ (compositum of fields).

(ii) If H is a factor group of G, then $K(H) \subseteq K(G)$.

(iii) $K(G)(G) = K(G)$.

Proof (i) The absolutely irreducible characters of the product group $G \times H$ have the form: $\psi(g,h) = \chi(g)\varphi(h)$ where χ and φ are are absolutely irreducible characters of G and H respectively.

(ii) Every representation of H may be factored through a representation of G.

(iii) Obvious from the definition of K(G).

Proposition 2.3.2

Let H be a subgroup of G such that two elements of H are conjugate in H if and only if they are conjugate in G. Then $K(H) \subseteq K(G)$.

Proof Let m be any multiple of the exponent of G, and hence a multiple of the exponent of H. Trivially $\Gamma(G) \subseteq \Gamma(H)$. The result follows by Theorem 2.1.1 and the definition of $\Gamma(G,K)$.

Theorem 2.3.3

Let H be a normal subgroup of prime index p in the finite group G, and let m be any multiple of the exponent of G. Then

(i) $\Gamma(G) \cap \Gamma(H)$ has index a power of p in $\Gamma(G)$,

(ii) if $K(G) = K$, then $[K(H):K] = p^n$ for some integer n.

Proof (i) If C is a class of G, then either $C \cap H = \emptyset$ or $C \subseteq H$ since H is a normal subgroup. Let $c \in C \subseteq H$, and write $C_G(c)$ for the centralizer of c in G, i.e. the set of $g \in G$ with $gc = cg$. Then $|C| = [G:C_G(c)]$. Restricting the natural map from G to G/H to the subgroup $C_G(c)$, we easily see that the image has order 1 or p, and that

the kernel is $C_G(c) \cap H = C_H(c)$. Hence

$$[H:C_H(c)] = [G:C_H(c)]/[G:H] = [G:C_G(c)][C_G(c):C_H(c)]/[G:H]$$
$$= |C|p^{-1}[C_G(c):C_H(c)] = \text{either } |C| \text{ or } p^{-1}|C|.$$

We deduce that either C is a conjugacy class in H or $C = C_1 \cup \ldots \cup C_p$ where the C_i are distinct conjugacy classes in H and $|C_i| = |C|/p$. In the latter case the group G/H acts on the set $C_1, \ldots, C_p$ by conjugation, and this action must be transitive since their union is a single conjugacy class in G. Hence there is a $g \in G$ such that $C_i = g^{-i} C_p g^i$ for $1 \leqslant i \leqslant p$, (after suitably reordering the C_i).

If C is a conjugacy class of G which remains a complete conjugacy class in H, then c^k is conjugate to c in H for all $c \in C$ and all $k \in \Gamma(G)$. If, on the other hand, C splits as above: $C = C_1 \cup \ldots \cup C_p$ with $C_i = g^{-i} C_p g^i$ a conjugacy class in H, let $c \in C_p$ and $k \in \Gamma(G)$. Then $c^k \in C_i$ for some i (since C is a single conjugacy class of G). Then c^k is conjugate in H to $g^{-i} c g^i$, and so $c^{k^2} = (c^k)^k$ is conjugate to $g^{-i} c^k g^i \in g^{-i} C_i g^i = C_j$ where $j \equiv 2i \bmod(p)$. Similarly $c^{k^3} = (c^{k^2})^k \in g^{-i} C_j g^i = C_{3i}$ where the subscript is read modulo p. In this way we show that $c^{k^p} \in C_p$ for all $k \in \Gamma(G)$ and all $c \in C_p$. C_p may be replaced in this argument by any other C_i by writing $C_j' = g^{-j} C_p' g^j$ where $C_p' = C_i = g^{-i} C_p g^i$.

This shows that $k^p \in \Gamma(H)$ for all $k \in \Gamma(G)$, and so that k^p is in $\Gamma(G) \cap \Gamma(H)$ for all $k \in \Gamma(G)$. (i) is now immediate from the structure theorems for finite abelian groups: $\Gamma(G) = P \oplus N$ where P is a p-group and N contains only elements of order prime to p. The image of the endomorphism $k \to k^p$ contains N, and so has index dividing $|P|$.

(ii) follows from (i), Theorem 2.1.1 and Galois theory.

Note The only case of this theorem which we shall need is when $p = 2$. Of course a subgroup of index 2 is automatically normal. The

simplest example of an application of this result is to the alternating group A_n which is a subgroup of index 2 in the symmetric group S_n. It is possible to show that $Q(S_n) = Q$, and hence we have $[Q(A_n):Q]$ = a power of 2.

Corollary 2.3.4

Let H be a normal subgroup of prime power index p^n in the finite group G. If $K(G) = K$, then $[K(H):K]$ = a power of p.

Proof By the elementary theory of p-groups we may interpolate a chain of subgroups: $G = H_1 \supset H_2 \supset \ldots \supset H_{n+1} = H$ with H_{i+1} normal in H_i and the successive quotients all cyclic of order p. Repeated applications of Theorem 2.3.3 eventually yield the result: $K(H_2) = K_2$ has degree a power of p over K, $K_3 = K_2(H_3)$ has degree a power of p over K_2, and we eventually find $K_n = K_{n-1}(H_n)$, $K_{n+1} = K_n(H) \supseteq K(H)$. The result follows since $[K(H):K]$ must divide $[K_{n+1}:K]$ which is a power of p.

Notes (i) The statement and proof of this corollary still hold if the word "normal" is replaced by "subnormal".

(ii) In fact the proof of Theorem 2.3.3 shows that the quotient group $\Gamma(G)/(\Gamma(G) \cap \Gamma(H))$ has exponent dividing p, and so is an elementary abelian p-group. Similarly the proof of the corollary could be extended to show that in this case the quotient has exponent dividing p^n.

2.4 Splitting fields, Schur indices and integral representations

In this section we collect together several results from character theory which give reasons for wanting to determine the fields K(G) for finite groups G. These results are quoted without proof and may be found in W. Feit:"Characters of Finite Groups", Benjamin (1967).

A field F is called a splitting field for the finite group G if every irreducible representation defined over F is absolutely irreducible. Let K be a field such that K(G) = K, and of course char(K) = 0.

Proposition 2.4.1

Let χ be an irreducible character of G, K a field of characteristic zero, and K(G) = K.

(i) There is a positive integer m such that $m\chi$ is the character of a representation defined over K. We define the Schur index $m_K(\chi)$ to be the smallest such positive integer.

(ii) If $m\chi$ is the character of some representation defined over K, then $m_K(\chi)$ divides m.

(iii) If F is a finite extension of K which is a splitting field for G, then $m_K(\chi)$ divides $[F:K]$.

(iv) If η is the character of some representation defined over K, then $m_K(\chi)$ divides (χ,η) where (,) is the usual inner product in the character ring of G.

(v) $m_K(\chi)$ divides $\chi(1)$.

Proof These are all special cases of results (11.2), (11.3), (11.4) and (11.5) in Feit's book.

Theorem 2.4.2

(i) If G is a group of exponent m, then $Q(\sqrt[m]{1})$ is a splitting field

for G.

(ii) Let $\{p_i\}$ be the set of primes dividing the order of G. Let $F \supseteq K$ be a field containing a primitive $(p_1 p_2 \ldots)^{th}$ root of unity, and also containing $\sqrt{-1}$ if $|G|$ is even. Then F is a splitting field for G.

(iii) Let p be a prime, and let $|G| = p^c b$ with $(p,b) = 1$. Let $F \supseteq K$ be a field containing a primitive b^{th} root of unity, and also containing either $\sqrt{-1}$ or $\sqrt{-3}$ if $p = 2$. Then F is a splitting field for G.

(iv) If $Q_p(G)$ is an unramified extension of Q_p, then there is an unramified extension F of Q_p which is a splitting field for G.

Proof (i) (16.3) in Feit's book - a theorem of Brauer.

(ii) (16.4) in Feit's book - a theorem of Solomon.

(iii) (16.5) in Feit's book - a theorem of Fong.

(iv) An easy consequence of (16.6) in Feit's book - another theorem of Brauer.

We quote one result from integral representation theory.

Theorem 2.4.3

Let D be a principal ideal domain with quotient field F. Every representation of G defined over F is similar to a representation defined over D.

Proof (4.1) in Feit's book.

A possible application of this arises when $Q_q(G) = Q_q$. $W(F_q)$ is a principal domain with quotient field Q_q. If we know that the character χ has $m_{Q_q}(\chi) = 1$, (for example by using Proposition2.4.1(iv)), then we know that χ arises from a representation defined over $W(F_q)$.

Section 3 Conjugacy Classes in the Classical Groups

3.1 Linear geometries and groups

Let V be a finite dimensional vector space over a field F and write GL(V) for the group of all F-linear automorphisms of V. If $\dim_F V = n$, then GL(V) is isomorphic to $GL_n(F)$, the multiplicative group of all invertible n X n matrices over F.

Let σ be a fixed automorphism of F satisfying σ^2 = identity, and write $\sigma(a) = \bar{a}$ for $a \in F$. (Thus $\bar{\bar{a}} = a$ for all a). Let $f: V \times V \longrightarrow F$ be a sesquilinear form on V, i.e. for all $x,y,u,v \in V$ and all $c,d \in F$

$$f(cx+dy,u) = cf(x,u) + df(y,u) \quad \text{and}$$

$$f(x,cu+dv) = \bar{c}f(x,u) + \bar{d}f(x,v).$$

We call the pair (V,f) a sesquilinear space over F.

Lemma 3.1.1

If (V,f) and (V,g) are both sesquilinear spaces over F such that $g(x,y) = 0$ whenever $f(x,y) = 0$, then there exists $\lambda \in F$ such that $g = \lambda f$. If $g \neq 0$ then λ is unique and non-zero.

Proof There is nothing to prove unless $g(x,y) \neq 0$ for some x and y, and in this case we must show that $g(x,y)/f(x,y) = g(u,v)/f(u,v)$ whenever $f(u,v) \neq 0$.

First note that if $f(u,v) \neq 0$ then there is a z in V satisfying $f(x,z) \neq 0 \neq f(u,z)$. (If neither z = y nor z = v satisfies this, then z = y+v must). Let $a = f(x,z) \neq 0 \neq f(u,z) = b$. Hence $f(bx-au,z) = ba - ab = 0$, and so $g(bx-au,z) = 0$ by hypothesis. Hence $g(x,z)/f(x,z) = g(u,z)/f(u,z)$ by the linearity of f and g in their first variable. Almost identical proofs yield the equations:

$$g(x,y)/f(x,y) = g(x,z)/f(x,z) \text{ and } g(u,z)/f(u,z) = g(u,v)/f(u,v).$$

Two sesquilinear spaces (V,f) and (W,g) over the same field F (with the same automorphism σ) are said to be conformally equivalent

Section 3 Conjugacy Classes in the Classical Groups

3.1 Linear geometries and groups

Let V be a finite dimensional vector space over a field F and write GL(V) for the group of all F-linear automorphisms of V. If $\dim_F V = n$, then GL(V) is isomorphic to $GL_n(F)$, the multiplicative group of all invertible n X n matrices over F.

Let σ be a fixed automorphism of F satisfying $\sigma^2 =$ identity, and write $\sigma(a) = \bar{a}$ for $a \in F$. (Thus $\bar{\bar{a}} = a$ for all a). Let $f: V \times V \longrightarrow F$ be a sesquilinear form on V, i.e. for all $x,y,u,v \in V$ and all $c,d \in F$

$$f(cx+dy,u) = cf(x,u) + df(y,u) \quad \text{and}$$

$$f(x,cu+dv) = \bar{c}f(x,u) + \bar{d}f(x,v).$$

We call the pair (V,f) a sesquilinear space over F.

Lemma 3.1.1

If (V,f) and (V,g) are both sesquilinear spaces over F such that $g(x,y) = 0$ whenever $f(x,y) = 0$, then there exists $\lambda \in F$ such that $g = \lambda f$. If $g \neq 0$ then λ is unique and non-zero.

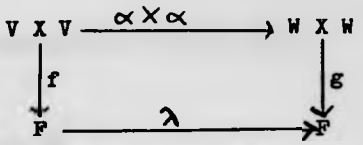
Proof There is nothing to prove unless $g(x,y) \neq 0$ for some x and y, and in this case we must show that $g(x,y)/f(x,y) = g(u,v)/f(u,v)$ whenever $f(u,v) \neq 0$.

First note that if $f(u,v) \neq 0$ then there is a z in V satisfying $f(x,z) \neq 0 \neq f(u,z)$. (If neither $z = y$ nor $z = v$ satisfies this, then $z = y+v$ must). Let $a = f(x,z) \neq 0 \neq f(u,z) = b$. Hence $f(bx-au,z) = ba - ab = 0$, and so $g(bx-au,z) = 0$ by hypothesis. Hence $g(x,z)/f(x,z) = g(u,z)/f(u,z)$ by the linearity of f and g in their first variable. Almost identical proofs yield the equations:

$$g(x,y)/f(x,y) = g(x,z)/f(x,z) \text{ and } g(u,z)/f(u,z) = g(u,v)/f(u,v).$$

Two sesquilinear spaces (V,f) and (W,g) over the same field F (with the same automorphism σ) are said to be conformally equivalent

if there is a vector space isomorphism $\alpha: V \longrightarrow W$ and a non-zero $\lambda \in F$ such that $g(\alpha x, \alpha y) = \lambda f(x,y)$ for all $x,y \in V$.



(i.e. the diagram on the right commutes).

The spaces are said to be equivalent if they are conformally equivalent with $\lambda = 1$. (Both conformal equivalence and equivalence are clearly equivalence relations).

We define the isometry group of (V,f)

$O(V,f)$ = the set of $t \in GL(V)$ with $f(tx,ty) = f(x,y)$ for all $x,y \in V$,

and the conformal isometry group

$CO(V,f)$ = the set of $t \in GL(V)$ such that $f(tx,ty) = 0$ whenever $f(x,y) = 0$.

The next proposition shows that both of these sets are subgroups of $GL(V)$.

Proposition 3.1.2

If $f \neq 0$, then

(i) for each $t \in CO(V,f)$ there is a unique $\lambda(t) \in F^*$ such that

$$f(tx,ty) = \lambda(t) f(x,y) \text{ for all } x,y \in V$$

(ii) $CO(V,f)$ = the set of $t \in GL(V)$ such that $f(tx,ty) = \kappa f(x,y)$ for some $\kappa \in F$ and all $x,y \in V$

(iii) There is an exact sequence of groups and homomorphisms:

$$1 \longrightarrow O(V,f) \hookrightarrow CO(V,f) \xrightarrow{\lambda} F^*$$

Proof (i) Apply Lemma 3.1.1 with $g(x,y) = f(tx,ty)$.

(ii) Immediate from (i) and the definition of $CO(V,f)$.

(iii) $CO(V,f)$ is a subgroup of $GL(V)$ by (ii). Clearly λ is a group homomorphism with kernel $O(V,f)$ by (i) and (ii).

Notes 1. If $f = 0$, then $O(V,f) = CO(V,f) = GL(V)$.

2. For $f \neq 0$ we define the multiplier of f by

$M(f)$ = image of λ = the set of all $\lambda(t)$ as t runs through $CO(V,f)$, a subgroup of F^*. If we replace F^* by $M(f)$ we get a short exact sequence in Proposition 3.1.2(iii).

3. If (V,f) and (W,g) are conformally equivalent with a vector space

isomorphism α as above, then we have a commutative diagram with exact rows:

$$\begin{array}{ccccccc} 1 & \longrightarrow & O(V,f) & \hookrightarrow & CO(V,f) & \longrightarrow & F^* \\ & & \downarrow \beta & & \downarrow \beta & & \downarrow \text{identity} \\ 1 & \longrightarrow & O(W,g) & \hookrightarrow & CO(W,g) & \longrightarrow & F^* \end{array}$$

where the vertical maps are all isomorphisms, and $\beta(t) = \alpha t \alpha^{-1}$ for $t \in CO(V,f)$.

A sesquilinear space (V,f) is said to be reflexive if $f(y,x) = 0$ whenever $f(x,y) = 0$.

Proposition 3.1.3

If (V,f) is a reflexive sesquilinear space, then either

(i) σ = identity, and $f(y,x) = -f(x,y)$ for all $x,y \in V$,

or (ii) f is conformally equivalent to a form h on V such that

$$h(y,x) = \overline{h(x,y)} \text{ for all } x,y \in V.$$

Proof Applying Lemma 3.1.1 with $g(x,y) = \overline{f(y,x)}$, we see that there is a constant $a \in F$ such that

$$f(y,x) = a.\overline{f(x,y)} \text{ for all } x \text{ and } y \in V$$

$$= a\bar{a}.f(y,x)$$ upon reversing the order of variables a second time. Unless f is identically zero, in which case there is nothing to prove, this shows that $a\bar{a} = 1$. If σ = identity, then $a^2 = 1$, so that $a = 1$ or -1 and in either case the proposition holds.

If $\sigma \neq$ identity, then $\bar{x}/x$ must take at least two values (including 1) for $x \in F^*$. Hence there is an $x \in F$ with $q = x + \overline{ax} \neq 0$. Also $aq = ax + a\overline{ax} = \bar{x} + ax = \bar{q}$. Define $h(x,y) = q.f(x,y)$, so that

$$h(y,x) = q.f(y,x) = qa.\overline{f(x,y)} = (qa/\bar{q}).\overline{h(x,y)}$$

$$= \overline{h(x,y)}.$$

It is customary to divide the conformal equivalence classes described in this proposition into three special cases:
(i) The unitary case: $\sigma \neq$ identity and $f(y,x) = \overline{f(x,y)}$ for all x,y.

f is called a hermitian form on V, and the groups, O(V,f) and CO(V,f), are called the unitary and conformal unitary groups of (V,f) respectively, and written U(V) and CU(V). (Explicit mention of the form f is rarely made if there is little danger of ambiguity).

(ii) The orthogonal case: σ = identity and f(x,y) = f(y,x) always. f is a symmetric bilinear form on V. The groups are called the orthogonal and conformal orthogonal groups of (V,f), and are written O(V) and CO(V) respectively.

(iii) The symplectic case: σ = identity and f(y,x) = -f(x,y) always. f is an antisymmetric form on V. The groups are called the symplectic and conformal symplectic groups of (V,f), and are written Sp(V) and CSp(V) respectively.

Note Cases (ii) and (iii) coincide if char(F) = 2. In this case it is customary to replace the antisymmetric condition on f by the stronger alternating condition: f(x,x) = 0 for all $x \in V$. (Of course, if char(F) $\neq$ 2, the anti/symmetric and alternating conditions are equivalent).

A sesquilinear space (V,f) is said to be non-degenerate if whenever f(x,y) = 0 for all $y \in V$, then x = 0. This is equivalent to the condition: "... if whenever f(x,y) = 0 for all $x \in V$, then y = 0", since it is not hard to see that both statements are equivalent to: "the matrix $(a_{ij}) = (f(e_i,e_j))$ is non-singular where $e_1,\ldots,e_n$ is any basis of V". (The equivalence of the two conditions is obvious directly in the reflexive case). A subspace, U, of V is called non-degenerate if the restriction of f to U gives a non-degenerate space (U,f).

Lemma 3.1.4

If U is a non-degenerate subspace of the sesquilinear space (V,f) and $g: U \longrightarrow F$ is an F-linear map, then there is a (unique) $u \in U$ such

$$g(x) = f(x,u) \text{ for all } x \in U.$$

Proof Consider the linear maps $g_i: U \longrightarrow F$ given by $g_i(x) = f(x,e_i)$ where $e_1,\ldots,e_n$ is a basis of U. These maps form a basis of the dual space of U, since otherwise there is a non-zero vector x in the kernel of all of them, and so $f(x,u) = 0$ for all $u \in U$, contradicting the non-degeneracy of U. Let $g = a_1g_1 + \ldots + a_ng_n$ with $a_i \in F$. Take $u = \bar{a}_1e_1 + \ldots + \bar{a}_ne_n$.

Let (V,f) be a reflexive sesquilinear space. Two subsets, X and Y of V are said to be orthogonal if $f(x,y) = 0$ whenever $x \in X$ and $y \in Y$. The relationship of orthogonality is clearly symmetric. A direct sum, $U_1 \oplus \ldots \oplus U_k$, of subspaces is called an orthogonal direct sum if U_i and U_j are orthogonal whenever $i \neq j$.

Proposition 3.1.5

If U is a non-degenerate subspace of the reflexive sesquilinear space (V,f), then $V = U \oplus U^{\perp}$, an orthogonal direct sum, where $U^{\perp}$ is the set of $x \in V$ such that $f(x,u) = 0$ for all $u \in U$.

Proof If $x \in U \cap U^{\perp}$, then $f(x,u) = 0$ for all $u \in U$, and so $x = 0$ since U is non-degenerate. Hence the sum is direct, and it is clearly orthogonal. ($U^{\perp}$ is certainly a subspace). Let $x \in V$. By Lemma 3.1.4 there is a $u \in U$ such that $f(y,x) = f(y,u)$ for all $y \in U$, and so $f(y,x-u) = 0$ for all $y \in U$. The equation $x = u + (x-u)$ shows that the direct sum equals the whole of V.

A basis $e_1,\ldots,e_n$ of a reflexive sesquilinear space (V,f) is called an orthogonal basis if $f(e_i,e_j) = 0$ whenever $i \neq j$.

Proposition 3.1.6

If (V,f) is a reflexive sesquilinear space and $f(x,y) + f(y,x) \neq 0$ for some $x,y \in V$, then V has an orthogonal basis.

Proof It is clearly enough to prove the result for some form conformally equivalent to f. So by Proposition 3.1.3 we may assume that $f(y,x) = \overline{f(x,y)}$ for all $x,y \in V$. By hypothesis there are $x,y \in V$

such that $f(x,y) = a$ with $a + \bar{a} \neq 0$. The identity

$$f(x+y,x+y) = f(x,x) + f(y,y) + f(x,y) + \overline{f(x,y)}$$

shows that $f(v,v) \neq 0$ for some $v \in V$. Let U be the non-degenerate subspace spanned by v, so that $V = U \oplus U^{\perp}$, an orthogonal direct sum. If the restriction of f to $U^{\perp}$ is identically zero, then any basis of $U^{\perp}$ extends v to an orthogonal basis of V. If f is not identically zero on $U^{\perp}$, there must exist $u,w \in U^{\perp}$ with $f(u,w) = a$, by the linearity of f in its first variable. In this case induction on the dimension of V completes the proof.

Proposition 3.1.7

If (V,f) is a symplectic space over F, then V has a basis $e_1,\ldots,e_n$, $f_1,\ldots,f_m$ such that $f(e_i,e_j) = f(f_i,f_j) = 0$ for all i and j, and

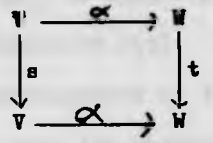
$$f(e_i,f_j) = \delta_{ij} \text{ (the Kronecker } \delta\text{)}.$$

Proof If f is identically zero let $e_1,\ldots,e_n$ be any basis of V. If $f(x,y) = a \neq 0$, let $e_1 = a^{-1}x$, and $f_1 = y$. Then let U be the non-degenerate subspace spanned by e_1 and f_1. By induction on the dimension of V, we may find a basis $e_2,\ldots,f_2,\ldots,f_n$ of $U^{\perp}$ with the required properties.

Note In this case (V,f) is non-degenerate if and only if $m = n$, and so $\dim_F V = 2n$ must be even for non-degeneracy.

3.2 Conjugacy classes in the general linear groups

Given a category $\mathcal{C}$ we may define a new category, Auto($\mathcal{C}$) whose objects are ordered pairs (V,s) with V an object of $\mathcal{C}$ and $s \in \mathrm{Aut}(V)$ = the group of invertible morphisms from V to itself. A morphism from (V,s) to (W,t) in Auto($\mathcal{C}$) is a morphism $\alpha: V \longrightarrow W$ of $\mathcal{C}$ such that the diagram on the right commutes. (It is easily checked that Auto($\mathcal{C}$) is a category with these definitions). Clearly two elements, s and t, of the group Aut(V) are conjugate if and only if (V,s) and (V,t) are isomorphic in Auto($\mathcal{C}$). We apply these remarks to the determination of conjugacy classes in various groups of linear transformations in the next few sections.



Firstly let $\mathcal{C}$ = VS, the category of finite dimensional vector spaces over a (fixed) field F, with F-linear maps as morphisms. We write Aut(V) = GL(V), the general linear group of V over F. Let x be transcendental over F, and let F[x] be the polynomial ring in x over F. Given $(V,s) \in \mathrm{Auto}(VS)$ we define a F[x]-module, V_s, by

additive group of V_s = V

$f(x)v = f(s)v$ for all $v \in V$ and all $f(x) \in F[x]$,

where s is considered as an element of the F-algebra of endomorphisms of V. The main results on conjugacy in GL(V) are summarised in the following theorem.

Theorem 3.2.1

(i) (V,s) and (W,t) are isomorphic in Auto(VS) if and only if V_s and W_t are isomorphic as F[x]-modules.

(ii) As F[x]-module, $V_s = \bigoplus_p V_s(p)$, where p(x) runs through the monic irreducible polynomials in F[x], and $V_s(p)$ is the subspace consisting of those $v \in V$ for which $p(s)^i v = 0$ for some $i > 0$. $V_s(p) \neq 0$ if and only if p(x) divides the characteristic

polynomial of s.

(iii) V_s and W_t are isomorphic $F[x]$-modules if and only if $V_s(p)$ and $W_t(p)$ are isomorphic for all monic irreducible $p(x) \in F[x]$.

(iv) $V_s(p) \cong (R/p^{n_1}R) \oplus \ldots \oplus (R/p^{n_k}R)$ where $R = F[x]$, $0 < n_1 \leq n_2 \leq \ldots \leq n_k$, and the sequence $(n_1, \ldots, n_k)$ is uniquely determined by s and the monic irreducible polynomial $p(x)$.

Proof (i) is obvious. (ii), (iii) and (iv) are consequences of the structure theorems for finitely generated modules over the prindipal ideal domain $F[x]$. For a detailed treatment see chapter XV of Lang's "Algebra".

An $F[x]$-module is said to be indecomposable if it cannot be written as an (internal) direct sum of two proper submodules.

Proposition 3.2.2

(i) $F[x]/p(x)^iF[x]$ is indecomposable if $p(x)$ is irreducible.

(ii) Two indecomposable $F[x]$-modules V_s and W_t are isomorphic if and only if (a) $\dim_F V = \dim_F W$, and

(b) s and t have the same eigenvalues (in some algebraic closure of F).

Proof (i) Suppose $F[x]/p(x)^iF[x] = M \oplus N$. Some $y \in M \cup N$ must be the image under the natural map of some $f(x) \in F[x]$ which is not divisible by $p(x)$. (Otherwise $M \oplus N$ would be contained in the proper submodule $p(x)F[x]/p(x)^iF[x]$). Since $p(x)$ is irreducible and $F[x]$ is a principal ideal domain, there are polynomials $h(x)$ and $k(x)$ such that

$$f(x)h(x) + p(x)^ik(x) = 1.$$

Hence the image, y, of $f(x)$ under the natural map generates the whole module, $F[x]/p(x)^iF[x]$, and so either M or N is the whole module.

(ii) The necessity of conditions (a) and (b) follows from elementary linear algebra.

Suppose that conditions (a) and (b) are satisfied. By (i)

above and Theorem 3.2.1(ii), (iii) and (iv)

$$V_s \cong F[x]/p(x)^i F[x] \text{ and } W_t \cong F[x]/q(x)^j F[x],$$

where p and q are monic irreducible polynomials, and i and j are positive integers. Since s and t have the same eigenvalues, $p = q$ by Theorem 3.2.1(ii). Also

$$i.\deg(p) = \dim_F V = \dim_F W = j.\deg(q) \text{ by (a)},$$

and so $i = j$.

Lemma 3.2.3

If $s \in GL(V)$, s^m = identity and $(k,m) = 1$, then s and s^k have the same invariant subspaces in V.

Proof If $sU \subseteq U$, then $s^k U \subseteq U$ trivially.

Conversely, since $(k,m) = 1$ there exists j with $jk \equiv 1 \bmod(m)$. If $s^k U \subseteq U$, then $sU = s^{jk}U = (s^k)^j U \subseteq U$.

We use these results to examine GL(V) for F a finite field.

Proposition 3.2.4

If $F = F_q$, the finite field of q elements, then

$$|GL(V)| = q^N(q-1)(q^2-1)\ldots(q^n-1) \text{ where } n = \dim_F V,\ N = n(n-1)/2.$$

Proof Let $e_1,\ldots,e_n$ be a basis of V. We count the possible choices of $t \in GL(V)$. Firstly te_1 may be chosen in q^n-1 ways to be a non-zero vector. Then te_2 may be chosen in q^n-q ways to be linearly independent of te_1. Proceeding thus we find that te_i may be chosen in q^n-q^{i-1} ways for $1 \leqslant i \leqslant n$. Multiplying all these numbers together gives the number of ways of choosing t, which is $|GL(V)|$.

Theorem 3.2.5

Let $F = F_q$, the finite field of q elements, and $\dim_F V = n$. Let S be the set of integers, k, such that s and s^k are conjugate for all $s \in GL(V)$. Then

(i) if $n = 1$, S is the set of integers, k, with $k \equiv 1 \bmod(q-1)$

(ii) if $n>1$, S is the set of integers, k, such that

$k \equiv q^{a_i} \bmod(q^i-1)$ for some integer a_i, for $1 \leqslant i \leqslant n$, and

$(k,q) = 1$.

Proof (i) $GL(V) = F_q^*$, a cyclic group of order q-1, by the theory of finite fields, and the result is obvious.

(ii) Consider the vector space $V = F[x]/g(x)F[x]$, and the endomorphism s induced by multiplication by x. s has minimal (and hence characteristic) polynomial $g(x)$, and so is singular if and only if $g(0) = 0$. This shows that every irreducible polynomial of degree less than or equal to n, except x, may arise as a factor of the characteristic polynomial of some $t \in GL(V)$.

In particular if $1 \leqslant i \leqslant n$, and c is a generator of the cyclic group $(F_{q^i})^*$, there is a $t \in GL(V)$ with eigenvalues c, c^q, ..., $c^{q^{i-1}}$ and 1 (repeated n-i times). Then t^k has eigenvalues c^k, c^{kq},... and 1. If $k \in S$, we must have $c^k = c^{q^a}$ for some integer a between 0 and i-1 by elementary linear algebra. (t and t^k must have the same eigenvalues). Since c has multiplicative order q^i-1, this shows that $k \equiv q^{a_i} \bmod(q^i-1)$ for some integer a_i, and this must hold for each i between 1 and n. If $k \in S$, k must certainly be prime to the order of GL(V), and so $(k,q) = 1$ by Proposition 3.2.4.

Conversely suppose that k satisfies the conditions in the statement of (ii) of this theorem, and let $s \in GL(V)$. No eigenvalue of s can have degree greater than n over F, and so, by the theory of the Frobenius automorphism for finite fields, s and s^k have the same eigenvalues on any s-invariant subspace of V. Also, by the conditions on k, $(k,q) = 1$ and $(k,q^i-1) = 1$ by Proposition 1.1.3(i), for $1 \leqslant i \leqslant n$. Hence $(k, |GL(V)|) = 1$ by Proposition 3.2.4 and repeated use of Proposition 1.1.1(ii)(b). In particular, k is prime to the order of s, and so s and s^k have the same invariant subspaces in V by Lemma 3.2.3.

By Theorem 3.2.1, V_s splits into a direct sum of s-invariant submodules: $V_s = M_1 \oplus \ldots \oplus M_j$, where each $M_i = F[x]/p(x)^j F[x]$ for some monic irreducible p and some integer j. Each M_i is indecomposable under the action of s by Proposition 3.2.2(i), and so is indecomposable under the action of s^k since they have the same invariant subspaces. By Proposition 3.2.2(ii) and the results in the previous paragraph, (M_i,s) and (M_i,s^k) are isomorphic in Auto(VS) for each i. Putting these isomorphisms together in the direct sum, we find that (V,s) and (V,s^k) are isomorphic in Auto(VS), as desired. Hence $k \in S$.

3.3 Conjugacy classes in the symplectic and orthogonal groups

Let F be a field and $\mathcal{E} = \pm 1$ be fixed. Let $\mathcal{C} = IPS$ be the category whose objects are finite dimensional vector spaces V over F on which is defined a non-degenerate F-bilinear form (,) such that $(v,u) = \mathcal{E}(u,v)$ for all u and v in V. A morphism $s:V \longrightarrow W$ in IPS is an F-linear map such that $(su,sv) = (u,v)$ for all u and v in V. (The same symbol (,) is used for the form on each V in IPS. This should not cause confusion). We use the same categorical construction as in the first paragraph of 3.2 to discuss the conjugacy classes in the groups Aut(V). If $\mathcal{E} = 1$, we may write Aut(V) = O(V), the orthogonal group of V, and if $\mathcal{E} = -1$ we write Aut(V) = Sp(V), the symplectic group of V. We note that an object, (V,s) of the category Auto(IPS) may be considered as an element of Auto(VS) by ignoring the bilinear form on V.

The results and proofs which follow (up to 3.3.6) are closely based on a paper of John Milnor: On Isometries of Inner Produot Spaces, (Inventiones Math., 8, (1969), p. 83-97). See also I.K. Cikunov: The Structure of isometric transformations of a symplectic or orthogonal vector space, (Soviet Math. Dokl. 6, (1965), p. 1479-81).

Let x be transcendental over F. Given a monic polynomial $g(x) = x^m + a_1 x^{m-1} + \ldots + a_m \in F[x]$ with $g(0) = a_m \neq 0$, we may define another such polynomial by

$$g^*(x) = g(0)^{-1} x^{\deg(g)} g(1/x) = (a_m x^m + a_{m-1} x^{m-1} + \ldots + 1)/a_m.$$

Lemma 3.3.1

(i) $\deg(g^*) = \deg(g)$ and $g^{**} = g$ whenever g^* is defined.

(ii) $(fg)^* = f^* g^*$ if the right-hand side is defined, where $fg(x) = f(x)g(x)$.

(iii) If g is monic irreducible, then so is g^*.

(iv) If $(V,s) \in$ Auto(IPS) and g(x) is monic with $g(0) \neq 0$, then

$$(g(s)u,v) = (u,g(s^{-1})v) = g(0)(u,s^{-\deg(g)} g^*(s)v) \text{ for } u,v \in V.$$

Proof (i) and (ii) are immediate from the definition of g^*, and (iii)

follows immediately from them. (iv) follows from the bilinearity of (,) and the identity

$$(su,v) = (s^{-1}su, s^{-1}v) = (u, s^{-1}v).$$

Lemma 3.3.2

If $(V,s) \in \text{Auto(IPS)}$ and $p(x)$, $q(x)$ are monic irreducible polynomials with $p^* \neq q$, then with the notation of Theorem 3.2.1(ii) $V_s(p)$ and $V_s(q)$ are orthogonal subspaces.

Proof If $V_s(p) \neq 0$ then $p(0) \neq 0$ since s is non-singular. Choose integers i and j such that $p(s)^i V_s(p) = q(s)^j V_s(q) = 0$, which is possible by Theorem 3.2.1, and choose polynomials $h(x)$ and $k(x)$ with $h(x)p^*(x)^i + k(x)q(x)^j = 1$, which is possible since $p^* \neq q$ and both are monic irreducible. In particular, $p^*(s)^i h(s) =$ identity on $V_s(q)$ and so $p^*(s)^i$ is invertible when restricted to $V_s(q)$. Hence for all $u \in V_s(p)$ and $v \in V_s(q)$ we have

$$0 = (0,v) = (p(s)^i u, v) = p(0)^i (u, s^{-i.\deg(p)} p^*(s)^i v).$$

Since $p(0) \neq 0$ and $s^{-i.\deg(p)} p^*(s)^i$ is a linear automorphism of $V_s(q)$, we must have $(u,v) = 0$ for all $u \in V_s(p)$ and $v \in V_s(q)$.

Theorem 3.3.3

Let $(V,s) \in \text{Auto(IPS)}$.

(i) V splits into an orthogonal direct sum of non-degenerate s-invariant subspaces: $V = \left[\bigoplus_{p=p^*} V_s(p)\right] \oplus \left[\bigoplus_{p \neq p^*} V_s(\{p,p^*\})\right]$ where p runs through the monic irreducible polynomials in $F[x]$, $V_s(p)$ is the set of $v \in V$ with $p(s)^i v = 0$ for some i, and $V_s(\{p,p^*\}) = V_s(p) \oplus V_s(p^*)$.

(ii) (V,s) and (W,t) are isomorphic in Auto(IPS) if and only if all the pairs: $(V_s(p),s)$ and $(W_t(p),t)$ for $p = p^*$, and $(V_s(\{p,p^*\}),s)$ and $(W_t(\{p,p^*\}),t)$ for $p \neq p^*$ are isomorphic in Auto(IPS).

(iii) For $p \neq p^*$, $(V_s(\{p,p^*\}),s)$ and $(W_t(\{p,p^*\}),t)$ are isomorphic in Auto (IPS) if and only if they are isomorphic in Auto(VS).

(iv) If $p = p^*$, $V_s(p)$ splits into an orthogonal direct sum of non-degenerate s-invariant subspaces:

$V_s(p) = V_s^1(p) \oplus \ldots \oplus V_s^m(p)$, where, in the notation of Theorem 3.2.1, $V_s^i(p)$ is a direct sum of modules of the form $F[x]/p(x)^iF[x]$.

Proof (i) There is a direct sum splitting of this form by Theorem 3.2.1. The splitting is orthogonal by the previous lemma. Each subspace in the splitting is non-degenerate since their orthogonal direct sum is.

(ii) Any isomorphism $h:(V,s) \longrightarrow (W,t)$ must map $V_s(p)$ to $W_t(p)$ by Theorem 3.2.1. The rest is obvious.

(iii) Given an isomorphism $h:(V_s(\{p,p^*\}),s) \longrightarrow (W_t(\{p,p^*\}),t)$ in Auto(VS), we know that h maps $V_s(p)$ isomorphically to $W_t(p)$ by Theorem 3.2.1. Choose a basis $e_1,\ldots,e_n$ of $V_s(p)$. Since $V_s(\{p,p^*\})$ is non-degenerate and $(e_i,e_j) = 0$ for all i and j by the last lemma, there is a unique basis $f_1,\ldots,f_n$ of $V_s(p^*)$ with $(e_i,f_j) = \delta_{ij}$ for all i and j. (This may be proved by the same method as Lemma 3.1.4). Similarly $he_1,\ldots,he_n$ is a basis of $W_t(p)$, and $W_t(p^*)$ has a unique basis $f_1',\ldots,f_n'$ with $(he_i,f_j') = \delta_{ij}$ for all i and j. Define a linear map $k:V_s(\{p,p^*\}) \longrightarrow W_t(\{p,p^*\})$ by

$$ke_i = he_i,\ kf_i = f_i' \text{ for } 1 \leqslant i \leqslant n.$$

Straightforward calculation shows that k is an isomorphism in Auto(IPS).

(iv) By Theorem 3.2.1 there is such a splitting in Auto(VS), but it need not be orthogonal. Let $V_s(p) = U \oplus V_s^m$ be such a linear splitting where V_s^m is a direct sum of modules of the form $F[x]/p(x)^mF[x]$, and $p(s)^{m-1}U = 0$.

Let $u \in V_s^m$ with $(v,u) = 0$ for all $v \in V_s^m$. Hence

$0 = (p(s)^{m-1}v,u)$ for all $v \in V_s(p) = U \oplus V_s^m$,

$= (v,p(0)s^{-\deg(p)}p(s)^{m-1}u)$ by Lemma 3.3.1. Hence $p(s)^{m-1}u = 0$ since $V_s(p)$ is non-degenerate, and so $u = p(s)u_1$ for some $u_1 \in V_s^m$. Similarly $(p(s)^{m-2}v,p(s)u_1) = 0$ for all $v \in V_s(p)$, and a repetition of the above argument shows that $p(s)^{m-1}u_1 = 0$, so that $u_1 = p(s)u_2$ for some $u_2 \in V_s^m$, and so $u = p(s)^2u_2$. Continuing in this

way, we eventually show that $u = p(s)^m u_m = 0$.

Thus V_s^m is a non-degenerate subspace of $V_s(p)$, and can be split off as an orthogonal direct summand by Proposition 3.1.5. $(V_s^m)^\perp$ is also clearly s-invariant, and applying the same argument to a direct sum decomposition of this subspace we eventually prove the existence of the desired decomposition.

We prove later that the isomorphism classes of the $(V_s^i(p),s)$ are uniquely determined by s and p in almost all cases.

Lemma 3.3.4

If $p(x)$ is monic irreducible and $p = p^*$, then

either (a) $p(x) = x + 1$ or $x - 1$,

or (b) $p(x) = x^k + a_1 x^{k-1} + \ldots + a_{k-1}x + 1$, where $k = 2d$ is even and $a_i = a_{k-i}$ for $1 \leqslant i \leqslant k-1$.

Proof If $p(a) = 0$, then $p(1/a) = 0$ by the definition of $p^* = p$. If $a = 1/a$ for some root a of p, then $a^2 = 1$ and we have case (a). Otherwise p must have an even number of roots and so have even degree. The rest follows by equating coefficients in the expressions for $p(x)$ and $p^*(x)$, noting that $p(1) \neq 0$.

In order to proceed further it is convenient to consider the cases (a) and (b) of this lemma separately. We deal with case (b) first. Define a rational function in $F(x)$ by

$$r(x) = x^{-d}p(x) = r(1/x).$$

Thus for $(V,s) \in \mathrm{Auto}(\mathrm{IPS})$, and $u,v \in V$, $(r(s)u,v) = (u,r(s)v)$ by Lemma 3.3.1(iv).

Write ξ for the image of x in the field $E = F[x]/p(x)F[x]$ under the natural map. Hence $E = F[\xi]$, and

$0 = p(\xi) = p^*(\xi) = p(0)^{-1}\xi^{\deg(p)}p(\xi^{-1})$, and so $p(\xi^{-1}) = 0$.

Hence there is a unique automorphism, $a \longmapsto \bar{a}$, of E fixing F such that

$\bar{\xi} = \xi^{-1}$. This automorphism is not the identity but $\bar{\bar{a}} = a$ for all $a \in E$. Keeping this notation we have:

Theorem 3.3.5

Let $p(x)$ be a separable irreducible monic polynomial over F of degree at least 2 with $p = p^*$. We use the notation of Theorem 3.3.3(iv). The vector space $H_s^i(p) = V_s^i(p)/p(s)V_s^i(p)$ over E admits a unique non-degenerate sesquilinear form $\langle (u),(v) \rangle$ over E satisfying

(i) $\langle (v),(u) \rangle = \varepsilon \overline{\langle (u),(v) \rangle}$

(ii) $\text{trace}_{E/F} \langle (u),(v) \rangle = (u, r(s)^{i-1} v)$ for $u,v \in V_s^i(p)$,

where (u) denotes the equivalence class of the vector u in the quotient space. The isomorphism class of $(V_s(p),s)$ in Auto(IPS) uniquely determines the equivalence classes of the sesquilinear spaces $H_s^i(p)$.

Note The hypothesis of separability is not necessary but this case is adequate for our purposes. See Milnor's paper for comments on this.

Proof Define a sequence of subspaces of $V_s(p)$ by $V(i) = \text{Ker}(p(s)^i)$, so that $0 = V(0) \subseteq V(1) \subseteq \ldots \subseteq V(m) = V_s(p)$. Clearly $V_s^i(p) \subseteq V(i)$ and it is easy to see that the inclusion map induces a vector space isomorphism: $H_s^i(p) = V_s^i(p)/p(s)V_s^i(p) \xrightarrow{\simeq} V(i)/[V(i-1)+p(s)V(i+1)] = H^i$. This isomorphism becomes E-linear when we define the action of E on both quotient spaces by $f(\xi)(u) = (f(s)u)$. We construct the inner product $\langle\ ,\ \rangle$ on H^i in a way that makes it clear that its equivalence class depends only on the isomorphism class of $(V_s(p),s)$ in Auto(IPS).

For $(u),(v) \in H^i$, define $((u),(v)) = (u, r(s)^{i-1} v) \in F$, which is well-defined since $r(s) = s^{-d} p(s)$. Next define

$\langle (u),(v) \rangle$ = the unique element $e' \in E$ such that $\text{trace}_{E/F}(ee') = (e(u),(v))$ for all $e \in E$. (This is well-defined since E is separable over F and $e \longmapsto (e(u),(v))$ is an F-linear map from E to F). Thus $\langle (u),(v) \rangle \in E$ is uniquely defined by the equation:

$$\text{trace}_{E/F}(e \langle (u),(v) \rangle) = (f(s)u, r(s)^{i-1} v) \text{ for all } e = f(\xi) \in E \quad (*)$$

Let $a = g(\xi)$, $e = f(\xi) \in E$. Then

$$\begin{aligned}\mathrm{trace}_{E/F}(e\langle a(u),(v)\rangle) &= \mathrm{trace}_{E/F}(e\langle (g(s)u),(v)\rangle) \\ &= (f(s)g(s)u, r(s)^{i-1}v) \text{ by } (*) \\ &= \mathrm{trace}_{E/F}(ea\langle (u),(v)\rangle) \text{ for all } e,a \in E.\end{aligned}$$

Hence $\langle a(u),(v)\rangle = a\langle (u),(v)\rangle$. Similarly we may show that $\langle (u)+(v),(w)\rangle = \langle (u),(w)\rangle + \langle (v),(w)\rangle$, and so $\langle\ ,\rangle$ is E-linear in its first variable. Also

$$\begin{aligned}\mathrm{trace}_{E/F}(e\langle (v),(u)\rangle) &= (f(s)v, r(s)^{i-1}u) = (r(s)^{i-1}v, f(s^{-1})u) \\ &= \varepsilon(f(s^{-1})u, r(s)^{i-1}v) \\ &= \varepsilon\,\mathrm{trace}_{E/F}(\bar{e}\overline{\langle (u),(v)\rangle}) \\ &= \mathrm{trace}_{E/F}(e\varepsilon\overline{\langle (u),(v)\rangle}) \text{ for all } e \in E,\end{aligned}$$

since $\varepsilon = \pm 1$. This proves property (i) of $\langle\ ,\ \rangle$ and also shows that it is sesquilinear. Property (ii) is the special case of (*) obtained by putting $e = 1$. If $\langle (u),(v)\rangle = 0$ for all $(u) \in H_s^i(p)$, then property (ii) implies that $r(s)^{i-1}v = 0$ since $(\ ,\ )$ is non-degenerate on $V_s(p)$, and hence $(v) = 0$ in $H_s^i(p)$. This proves that $\langle\ ,\ \rangle$ is non-degenerate.

Conversely suppose we are given a sesquilinear form $\langle\ ,\ \rangle$ on H^i satisfying (ii). Let $e = f(\xi) \in E$, $(u),(v) \in H^i$. Then

$$\begin{aligned}\mathrm{trace}_{E/F}(e\langle (u),(v)\rangle) &= \mathrm{trace}_{E/F}\langle e(u),(v)\rangle \text{ by sesquilinearity} \\ &= (f(s)u, r(s)^{i-1}v) \text{ by property (ii).}\end{aligned}$$

So $\langle\ ,\ \rangle$ necessarily satisfies equation (*) which determines a unique sesquilinear form on H^i.

Corollary 3.3.6

The equivalence classes of the sesquilinear spaces, $H_s^i(p)$ for $1 \leq i \leq m$, uniquely determine the isomorphism class of $(V_s(p), s)$ in Auto(IPS) for $\mathrm{char}(F) \neq 2$.

Note The restriction on the characteristic of F is not needed for the result, but is needed for the proof presented below. See Milnor's paper for comments on this. However the same restriction on the

characteristic is needed for some later results.

Proof First we consider the case $\varepsilon = 1$, (the orthogonal case). By Proposition 3.1.6 $H_s^i(p)$ has an orthogonal basis $(u_1),\ldots,(u_r)$ with $\langle (u_j),(u_j)\rangle = c_j = \bar{c}_j \neq 0$, since $\langle\ ,\ \rangle$ is non-degenerate. We choose a coset representative $u_{(j)}$ of (u_j) in $V_s^i(p)$ so that we may describe the inner product $(\ ,\ )$ as simply as possible.

First choose any representative u_1 of (u_1), and set $u_1' = (1 + a(s)r(s))u_1$ where the polynomial $a(x)\in F[x]$ is to be chosen later. For $0 \leqslant k < d$, where $2d = \deg(p)$, we have

$(s^k r(s)u_1', u_1') = (s^k r(s)^{i-2}u_1, u_1) + R_k$, where

$R_k = (s^k r(s)^{i-2}u_1, a(s)r(s)u_1) + (s^k r(s)^{i-1}a(s)u_1, u_1 + a(s)r(s)u_1)$

$= (s^k r(s)^{i-1}a(s^{-1})u_1, u_1) + (s^k r(s)^{i-1}a(s)u_1, u_1) + 0$ by Lemma 3.3.1 and since $r(s)^i = 0$ on $V_s^i(p)$

$= (s^k(a(s)+a(s^{-1}))u_1, r(s)^{i-1}u_1)$ by bilinearity and Lemma 3.3.1

$= \mathrm{trace}_{E/F}(\xi^k(a(\xi)+a(\bar{\xi}))\langle (u_1),(u_1)\rangle)$ by equation (*) in the last proof

$= \mathrm{trace}_{K/F}((\xi^k + \bar{\xi}^k)c_1(a(\xi)+a(\bar{\xi})))$ since $c_1 \in K$ and $a(x)\in F[x]$,

where K is the subfield of elements of E fixed under the involution $a \longmapsto \bar{a}$. (We have used the transitivity of the trace map for $F \subseteq K \subseteq E$).

The set of $(\xi^k + \bar{\xi}^k)c_1$ for $0 \leqslant k < d$ is a basis of K over F, and so there is a unique $h \in K$ such that

$$\mathrm{trace}_{K/F}((\xi^k + \bar{\xi}^k)c_1 h) = -(s^k r(s)^{i-2}u_1, u_1) \text{ for } 0 \leqslant k < d.$$

$h = e + \bar{e}$ for some $e \in E$ since $\mathrm{trace}: E \longrightarrow K$ is surjective. We choose $a(x)\in F[x]$ such that $a(\xi) = e \in E = F[\xi]$. With this choice

$$(s^k r(s)^{i-2}u_1', u_1') = 0 \text{ for } 0 \leqslant k < d.$$

Next we let $u_1'' = (1 + b(s)r(s)^2)u_1'$, and an almost identical calculation shows that we may choose $b(x)\in F[x]$ such that

$$(s^k r(s)^j u_1'', u_1'') = 0 \text{ for } 0 \leqslant k < d \text{ and } j = i-2,\ i-3.$$

Proceeding in this fashion we eventually construct $u_{(1)}$ such that

$(s^k r(s)^j u_{(1)}, u_{(1)}) = 0$ for $0 \leqslant k < d$ and $0 \leqslant j \leqslant i-2$

$= 0$ for $0 \leqslant k < d$ and $0 \leqslant j \neq i-1$, since

$r(s)^i V_s^i(p) = 0$. Also

$$(s^k r(s)^{i-1} u_{(1)}, u_{(1)}) = (s^k u_{(1)}, r(s)^{i-1} u_{(1)})$$
$$= \text{trace}_{E/F}(\xi^k c_1) \text{ for all } k, \text{ by } (*),$$

(it is clear from the construction that $(u_{(1)}) = (u_1)$ in $H_s^i(p)$).

Next suppose $k \geqslant d$. Then

$$(s^k r(s)^j u_{(1)}, u_{(1)}) = (u_{(1)}, s^{-k}(r(s)^j) u_{(1)}) = (s^{-k} r(s)^j u_{(1)}, u_{(1)})$$
$$= \tfrac{1}{2}((s^k + s^{-k}) r(s)^j u_{(1)}, u_{(1)}).$$

(This is the first use made of the condition: $\text{char}(F) \neq 2$). $x^k + x^{-k}$ and $r(x)$ can be written as polynomials in the variable $x + x^{-1}$ of degrees k and d respectively, (remembering that $r(x) = r(1/x)$). Hence $x^k + x^{-k} = t(x+x^{-1}) r(x) + q(x+x^{-1})$ with $\deg(t) < k$, $\deg(q) < d$. So

$$((s^k + s^{-k}) r(s)^j u_{(1)}, u_{(1)}) = (t(s+s^{-1}) r(s)^{j+1} u_{(1)}, u_{(1)})$$
$$+ (q(s+s^{-1}) r(s)^j u_{(1)}, u_{(1)}),$$

and so these inner products may be found by induction on k. Note that all these inner products $(s^k r(s)^j u_{(1)}, u_{(1)})$ are determined by k, j and c_1 only.

A basis of the vector subspace $F[s]u_{(1)}$ of $V_s^i(p)$ is given by the set of $s^k r(s)^j u_{(1)}$ for $0 \leqslant j \leqslant i-1$ and $0 \leqslant k \leqslant 2d-1$, and the inner product of two basis elements is given by

$$(s^k r(s)^j u_{(1)}, s^{k'} r(s)^{j'} u_{(1)}) = (s^{k-k'} r(s)^{j+j'} u_{(1)}, u_{(1)})$$
$$= (s^{k'-k} r(s)^{j+j'} u_{(1)}, u_{(1)}),$$

and so depends only on $j+j'$, $|k-k'|$, and c_1.

Suppose that $\sum_{\substack{0 \leqslant k < 2d \\ 0 \leqslant j < i}} a_{kj} s^k r(s)^j u_{(1)}$ is orthogonal to every vector in the subspace $F[s]u_{(1)}$, for some $a_{kj} \in F$. Taking its inner product with $s^{-k'} r(s)^{i-1} u_{(1)}$, remembering that $r(s)^i u_{(1)} = 0$, and using equation (*) from the proof of Theorem 3.3.5, we find that $0 = \text{trace}_{E/F}(\sum_k a_{k0} \xi^{k+k'} c_1) = \text{trace}_{E/F}(a_0 \xi^{k'} c_1)$ for all k', where $a_0 = \sum_{0 \leqslant k < 2d} a_{k0} \xi^k$. $a_0 = 0$, since the set of $\xi^{k'} c_1$ spans E over F as k' runs through the positive integers. Hence each $a_{k0} = 0$ since the set of ξ^k for $0 \leqslant k < 2d$ are linearly independent over F.

Similarly, taking an inner product with $s^{-k'}r(s)^{i-2}u_{(1)}$, we may show that each $a_{kl} = 0$. Proceeding in this fashion we show that $F[s]u_{(1)}$ is a non-degenerate subspace of $V_s^i(p)$.

Hence $V_s^i(p) = F[s]u_{(1)} \oplus U$, an orthogonal direct sum. Copying the computation above to find $u_{(1)}$, we start by choosing any representative u_2 of (u_2) from U, and find a representative $u_{(2)}$ in U such that the inner product on $F[s]u_{(2)}$ is determined by c_2, and is non-degenerate. It is clear from this construction that the isomorphism class of $(V_s^i(p),s)$ is uhiquely determined by the numbers $c_1,\ldots,c_r \in K$, and so is certainly determined by the equivalence class of the sesquilinear space $H_s^i(p)$. This completes the proof.

The case $\varepsilon = -1$ can be reduced to the case $\varepsilon = 1$ as follows: define a new inner product on $V_s(p)$ by

$$u.v = ((s-s^{-1})u,v) = v.u$$

This has the effect of multiplying all the associated forms $\langle , \rangle$ by $\xi - \bar{\xi}$, and so all the appropriate results for (,) can be deduced from those for the new inner product.

Corollary 3.3.7

In the notation of Theorem 3.3.3(iv), the isomorphism class of $(V_s(p),s)$ in Auto(IPS) uniquely determines and is determined by the isomorphism classes of the $(V_s^i(p),s)$ for $1 \leqslant i \leqslant m$.

Now we return to case (a) of Lemma 3.3.4.

Theorem 3.3.8

Let $p(x) = x + 1$ or $x - 1$, and assume that $\mathrm{char}(F) \neq 2$. We use the notation of Theorem 3.3.3. The vector space $U_{\bar{s}}^i(p) = V_{\bar{s}}^i(p)/p(s)V_{\bar{s}}^i(p)$ over F admits a non-degenerate bilinear form $\langle\langle (u),(v) \rangle\rangle$ defined by

$$\langle\langle (u),(v) \rangle\rangle = ((s-s^{-1})^{i-1}u,v) .$$

and satisfying $\langle\langle (v),(u) \rangle\rangle = (-1)^{i-1}\varepsilon \langle\langle (u),(v) \rangle\rangle$. The equivalence classes of the bilinear spaces $U_s^i(p)$ for $1 \leqslant i \leqslant m$ uniquely determine

the isomorphism class of $(V_s(p),s)$ in Auto(IPS).

Note The restriction on the characteristic of F is essential for this theorem.

Proof $\langle\ ,\ \rangle$ is well-defined since $(s-s^{-1})^i = 0$ on $V_s^i(p)$. It is non-degenerate since $(\ ,\ )$ is, and satisfies the symmetry (or antisymmetry) condition since $((s-s^{-1})u,v) = (u,(s^{-1}-s)v)$ for all u and $v \in V$. (The whole proof of this theorem resembles a very easy case of Theorem 3.3.5 and Corollaries 3.3.6 and 3.3.7 all combined).

Next we show that the equivalence class of the bilinear space $U_s^i(p)$ determines the isomorphism class of $(V_s^i(p),s)$. (It may be shown that the equivalence class is determined by the isomorphism class as at the beginning of the proof of Theorem 3.3.5). We deal with the case $\varepsilon = 1$. (The proofs for the case $\varepsilon = -1$ are almost identical).

Let $D = s-s^{-1}$. Then $(Du,v) = -(u,Dv)$ for all $u,v \in V$, and $\langle(u),(v)\rangle = (D^{i-1}u,v)$. Suppose that i is odd. Then $\langle\ ,\ \rangle$ is symmetric. Choose an orthogonal basis $(u_1),\ldots,(u_r)$ of $U_s^i(p)$ with $\langle(u_j),(u_j)\rangle = c_j \neq 0$. Choose an arbitrary representative u_1 of the coset (u_1). Since $(\ ,\ )$ is symmetric, $(D^k u_1,u_1) = 0$ automatically if k is odd. Write $u_1' = u_1 + aD^2u_1$. By a much simplified version of the argument in the proof of Corollary 3.3.6 we find that a may be chosen in F so that $(D^{i-3}u_1',u_1') = 0$. (i is odd). Adding a suitable multiple of D^4u_1' to u_1' and repeating the argument, we eventually find that u_1 could have beem chosen so that $(D^ju_1,u_1) = 0$ for $0 \leqslant j \neq i-1$ and $(D^{i-1}u_1,u_1) = c_1$. Using the set of D^ju_1 for $0 \leq j \leq i-1$ as a basis of $F[s]u_1$ we may imitate the rest of the proof of Corollary 3.3.6 after showing that the subspace is non-degenerate.

If i is even, then $\langle\ ,\ \rangle$ is antisymmetric, and by Proposition 3.1.7 $U_s^i(p)$ has a basis of vectors $(e_i),(f_j)$ with $\langle\langle(e_i),(f_j)\rangle\rangle = \delta_{ij}$. By the method of the last paragraph we may choose a representative e_1 of (e_1) such that $(D^je_1,e_1) = 0$ for all j, by modifying the first choice with multiples of both De_1 and Df_1 etc. Then f_1 is chosen similarly

so that $(D^j e_1, e_1) = 0 = (D^j f_1, f_1)$ for all j, $(D^j e_1, f_1) = 0$ unless $j = i-1$, $(D^{i-1} e_1, f_1) = 1$, and e_1, f_1 lie in the subspace spanned by (e_1) and (f_1). The proof then continues in the usual fashion after showing that this subspace is non-degenerate.

The study of these groups over a finite field and the proof of the theorem corresponding to Theorem 3.2.5 for these groups is postponed to a later section.

3.4 Conjugacy classes in the conformal symplectic and orthogonal groups

Let F be a field and $\epsilon = \pm 1$ be fixed. Let $\mathcal{C} = CIPS$ be the category whose objects are finite dimensional vector spaces V over F on which is defined a non-degenerate F-bilinear form $(\ , \ )$ such that $(v,u) = \epsilon (u,v)$ for all u and v in V. A morphism $s:V \longrightarrow W$ in CIPS is an F-linear map such that $(su,sv) = 0$ if and only if $(u,v)=0$ for $u,v \in V$. (The same symbol $(\ , \ )$ is used for the bilinear forms on V and W. This should not cause confusion). We use the same categorical construction as in the first paragraph of 3.2 to discuss the conjugacy classes in the groups Aut(V). If $\epsilon = 1$, we may write Aut(V) = CO(V), the conformal orthogonal group of V, and if $\epsilon = -1$ we write Aut(V) = CSp(V), the conformal symplectic group of V. We note that an object (V,s) of the category Auto(CIPS) may be considered as an element of Auto(VS) by ignoring the bilinear form on V. We also note that the categories, IPS and Auto(IPS), of 3.3 are both subcategories of the corresponding categories of this section. The treatment of conjugacy classes given below is very similar to the treatment in 3.3.

Let $s:V \longrightarrow W$ be a morphism in CIPS. Applying Lemma 3.1.1 with $f(u,v) = (u,v)$ and $g(u,v) = (su,sv)$, we see that there is a unique $\lambda(s) \in F^*$ such that $(su,sv) = \lambda(s)(u,v)$ for all $u,v \in V$. If s and t are morphisms in CIPS whose composition st is defined, then clearly $\lambda(st) = \lambda(s)\lambda(t)$. (Note that this use of the map λ is consistent with the definition of it given in Proposition 3.1.2(i)). From this and the definition of morphisms in Auto(CIPS), it follows that two objects (V,s) and (W,t) are isomorphic in Auto(CIPS) only if $\lambda(s) = \lambda(t)$. From now on we assume that all objects (V,s) under consideration have the same value of $\lambda(s) = \lambda$.

Let x be transcendental over F. Given a monic polynomial $g(x) = x^m + a_1 x^{m-1} + \ldots + a_m \in F[x]$ with $g(0) = a_m \neq 0$, we may define another such polynomial by

$$g^*(x) = g(0)^{-1}x^{\deg(g)}g(\lambda/x)$$
$$= (a_m x^m + a_{m-1}\lambda x^{m-1} + \ldots + a_i \lambda^{m-i}x^i + \ldots + \lambda^m)/a_m.$$

(This generalises the situation in 3.3 where $\lambda = 1$ throughout).

With minor modifications many of the results of 3.3 still hold.

Lemma 3.4.1

(i) $\deg(g^*) = \deg(g)$ and $g^{**} = g$ whenever g^* is defined.

(ii) $(fg)^* = f^*g^*$ if the right-hand side is defined, where $fg(x)=f(x)g(x)$.

(iii) If g is monic irreducible, then so is g^*.

(iv) If $(V,s)\in$ Auto(CIPS) with $\lambda(s) = \lambda$ and $g(x)$ is monic with $g(0) \neq 0$, then for $u,v\in V$

$$(g(s)u,v) = (u,g(\lambda s^{-1})v) = g(0)(u,s^{-\deg(g)}g^*(s)v)$$

Proof (i) and (ii) are immediate from the definition of g^*, and (iii) follows immediately from them. (iv) follows from the bilinearity of $(\ ,\ )$ and the identity

$$(su,v) = \lambda(s^{-1}su,s^{-1}v) = (u,\lambda s^{-1}v).$$

Lemma 3.4.2

If $(V,s)\in$ Auto(CIPS) and $p(x)$, $q(x)$ are monic irreducible polynomials with $p^* \neq q$, then with the notation of Theorem 3.2.1(ii), $V_s(p)$ and $V_s(q)$ are orthogonal subspaces.

Proof Almost identical to the proof of Lemma 3.3.2, but using Lemma 3.4.1 instead of Lemma 3.3.1 towards the end of the proof.

Theorem 3.4.3

Let $(V,s)\in$ Auto(CIPS).

(i) V splits into an orthogonal direct sum of non-degenerate s-invariant subspaces: $V = \left[\bigoplus_{p=p^*} V_s(p)\right]\oplus\left[\bigoplus_{p\neq p^*} V_s(\{p,p^*\})\right]$ where p runs through the monic irreducible polynomials in $F[x]$, $V_s(p)$ is the set of $v\in V$ with $p(s)^i v = 0$ for some i, and $V_s(\{p,p^*\}) = V_s(p)\oplus V_s(p^*)$.

(ii) (V,s) and (W,t) are isomorphic in Auto(CIPS) if and only if

there are isomorphisms in Auto(CIPS):

$i_p:(V_s(p),s) \longrightarrow (W_t(p),t)$ for all p with $p = p^*$, and

$i_p:(V_s(\{p,p^*\}),s) \longrightarrow (W_t(\{p,p^*\}),t)$ for all p with $p \neq p^*$,

such that $\lambda(i_p)$ is the same for all p.

(iii) For $p \neq p^*$, there is an isomorphism $i:(V_s(\{p,p^*\}),s) \rightarrow (W_t(\{p,p^*\}),t)$ with $\lambda(i) = K \neq 0$ if and only if the two objects are isomorphic in Auto(VS).

(iv) If $p = p^*$, $V_s(p)$ splits into an orthogonal direct sum of non-degenerate s-invariant subspaces:

$$V_s(p) = V_s^n(p) \oplus \ . \ . \ . \oplus V_s^m(p),$$

where in the notation of Theorem 3.2.1, $V_s^i(p)$ is a direct sum of submodules isomorphic to modules of the form $F[x]/p(x)^iF[x]$.

Proof (i) As for Theorem 3.3.3(i)

(ii) The necessity of the conditions follows as in the proof of Theorem 3.3.3(ii), noting that if $i:(V,s) \longrightarrow (W,t)$ is an isomorphism in Auto(CIPS), $\lambda(i) = \lambda(i_p)$ where i_p is the restriction of i to the appropriate non-degenerate subspace. Conversely we can easily put all the isomorphisms i_p together to obtain an isomorphism for the direct sum.

(iii) Define the basis elements

$$e_i,\ f_j,\ he_i, \text{ and } f_k'$$

as in the proof of Theorem 3.3.3 (iii). Then define a linear isomorphism:

$$i:(V_s(\{p,p^*\}),s) \longrightarrow (W_t(\{p,p^*\}),t)$$

by

$k(e_i) = h(e_i)$ for $1 \leqslant i \leqslant n$, and

$k(f_i) = K(f_i')$ for $1 \leqslant i \leqslant n$.

A straightforward calculation shows that k is an isomorphism in Auto(CIPS) with $\lambda(k) = K$.

(iv) With a few notational changes, the proof of Theorem 3.3.3(iv) still works.

The result analogous to Lemma 3.3.4 is a little more complicated for the category CIPS than for IPS:

Lemma 3.4.4

If $p(x)$ is monic irreducible with $p = p^*$, then one of the following conditions must hold:

(a) $p(x) = x - a$ where $a^2 = \lambda$,

or (b) $p(x) = x^2 - \lambda$ where λ is not a square in F,

or (c) $p(x) = x^k + a_1 x^{k-1} + \ldots + a_{k-1} x + \lambda^d$,
where $\deg(p) = k = 2d$ is even and $a_i = \lambda^{i-d} a_{k-i}$
for $1 \leq i \leq k-1$.

Proof If $p(a) = 0$, then $p(\lambda/a) = 0$ also by the definition of $p^* = p$. If $a = \lambda/a$ for some root a of p, then $a^2 = \lambda$, and we must have either (a) or (b) depending on whether $a \in F$ or not.

Otherwise p must have an even number of roots, and so

$$p(x) = x^k + a_1 x^{k-1} + \ldots + a_{k-1} x + a_k,$$

where $k = 2d$ is even. For $1 \leq i \leq k-1$ the coefficient of x^{k-i} in $p = p^*$ is $a_i = a_{k-i} \lambda^i / a_k$, and the constant term definition of $p^*(x)$. Hence $a_k^2 = \lambda^k = \lambda^{2d}$.

If $a_k = \lambda^d$, we obtain case (c) of the lemma. If $a_k = -\lambda^d$, it is easily verified that $p(a) = 0$, where a is any square root of λ, and so we are reduced to the case in the first paragraph of this proof.

Note For fixed λ, only one of cases (a) and (b) may arise. (b) is a special case of (c) if $\mathrm{char}(F) = 2$. If we set $\lambda = 1$, we just recover Lemma 3.3.4.

In order to proceed further it is convenient to consider the cases (a), (b) and (c) of this lemma separately. We deal with case (a) first. Define a rational function in F(x) by

$$\begin{aligned} r(x) &= x^{-d}p(x) = x^{-d}p^*(x) = x^{-d}p(0)^{-1}x^{2d}p(\lambda/x) \\ &= x^d\lambda^{-d}p(\lambda/x) = (\lambda/x)^{-d}p(\lambda/x) \\ &= r(\lambda/x). \end{aligned}$$

Thus for $(V,s)\in$ Auto(CIPS) with $\lambda(s) = \lambda$, and $u,v\in V$,

$(r(s)u,v) = (u,r(s)v)$ by Lemma 3.4.1(iv).

Write ξ for the image of x in the field $E = F[x]/p(x)F[x]$ under the natural map. Hence $E = F[\xi]$, and

$0 = p(\xi) = p^*(\xi) = p(0)^{-1}\xi^{\deg(p)}p(\lambda/\xi)$, and so $p(\lambda/\xi) = 0$. Hence there is a unique automorphism, $a\mapsto\bar{a}$, of E fixing F pointwise such that $\bar{\xi} = \lambda/\xi$. This automorphism is not the identity, (except perhaps in characteristic 2 when (b) and (c) coincide), but $\bar{\bar{a}} = a$ for all $a\in E$. Keeping this notation we have:

Theorem 3.4.5

Let p(x) be a separable irreducible monic polynomial over F satisfying condition (c) of Lemma 3.4.4. We use the notation of Theorem 3.4.3(iv). The vector space $H_s^i(p) = V_s^i(p)/p(s)V_s^i(p)$ over E admits a unique non-degenerate sesquilinear form $\langle(u),(v)\rangle$ over E satisfying

(i) $\langle(v),(u)\rangle = \varepsilon\,\overline{\langle(u),(v)\rangle}$

(ii) $\mathrm{trace}_{E/F}\langle(u),(v)\rangle = (u,r(s)^{i-1}v)$ for $u,v\in V_s^i(p)$,

where (u) denotes the equivalence class of the vector u in the quotient space. The isomorphism class of $(V_s(p),s)$ in Auto(CIPS) determines the equivalence class of the sesquilinear space $H_s^i(p)$ up to multiplication of $\langle\ ,\ \rangle$ by elements of F. (The same element of F must be used for each i).

Proof The proof of Theorem 3.3.5 goes through with a few minor modifications. The last two sentences are easy to prove.

Corollary 3.4.6

Suppose that char(F) $\neq 2$, and that we have conformal equivalences $\alpha_i : H_s^i(p) \longrightarrow H_t^i(p)$ for all i such that $\langle \alpha_i(u), \alpha_i(v) \rangle = K\langle\langle (u),(v) \rangle\rangle$ for all appropriate (u),(v) and i with K a fixed non-zero element of F. Then there is an isomorphism $w:(V_s(p),s) \longrightarrow (W_t(p),t)$ in Auto(CIPS) with $\lambda(w) = K$.

Proof The method used to prove Corollary 3.3.6 in the orthogonal case, ($\varepsilon = 1$), gives the desired result here. We must replace the polynomials in the variable $x+x^{-1}$ by polynomials in the variable $x + \lambda x^{-1}$. For each $V_s^i(p)$ we may construct a basis with respect to which the matrix of inner products depends linearly on the constants $c_1,\ldots,c_r$ derived from an orthogonal basis of $H_s^i(p)$. ("linearly" means "F-linearly" here). Similarly we may construct a basis of $W_t^i(p)$ with respect to which the matrix of inner products depends linearly on $Kc_1,\ldots,Kc_r$. The obvious mapping defined on these bases has all the required properties.

The reduction of the symplectic case, ($\varepsilon = -1$), to the orthogonal case, ($\varepsilon = 1$), may be achieved by defining a new symmetric inner product on $V_s(p)$ by

$$u.v = ((s-\lambda s^{-1})u,v),$$

and continuing as for Corollary 3.3.6.

Now we return to the cases (a) and (b) of Lemma 3.4.4, and find that both may be treated by the methods of Theorem 3.3.8.

Theorem 3.4.7

Let p(x) be a monic irreducible polynomial satisfying (a) of Lemma 3.4.4. We use the notation of Theorem 3.4.3(iv). Let $D = s - \lambda s^{-1}$. The vector space $U_s^i(p) = V_s^i(p)/p(s)V_s^i(p)$ over F admits a non-degenerate bilinear form $\langle (u),(v) \rangle$ defined by

$$\langle (u),(v) \rangle = (D^{i-1}u,v),$$

and satisfying $\langle (v),(u)\rangle = (-1)^{i-1}\varepsilon \langle (u),(v)\rangle$. (We assume that char(F) $\neq 2$). If we are given conformal equivalences $\alpha_i : U_s^i(p) \longrightarrow U_t^i(p)$ for all i such that $\langle \alpha_i(u), \alpha_i(v)\rangle = \kappa \langle (u),(v)\rangle$ for all appropriate (u),(v) and i with κ a fixed non-zero element of F, then there is an isomorphism $w:(V_s(p),s) \longrightarrow (W_t(p),t)$ in Auto(CIPS) with $\lambda(w) = \kappa$.

Proof If p(x) satisfies (a) of Lemma 3.4.4, the method of proof of Theorem 3.3.8, (which is itself an easy version of the proof of Theorem 3.3.5 and Corollaries 3.3.6 and 3.3.7), may be modified in an obvious way to yield the desired result.

Case (b) of Lemma 3.4.4 shares characteristics of both (a) and (c), and may be treated by a mixture of the methods used for them. Let $E = F[x]/(x^2-\lambda)F[x]$, where λ is not a square in F, and let ξ be the image of x in E under the natural map. E has an automorphism , $a \longmapsto \bar{a}$, fixing F pointwise and such that $\bar{\xi} = -\xi$. With this notation we have:

Theorem 3.4.8

Let p(x) be monic irreducible and satisfy (b) of Lemma 3.4.4. We use the notation of Theorem 3.4.3(iv). Let $D = s^{-1}p(s) = s - \lambda s^{-1}$. The vector space $H_s^i(p) = V_s^i(p)/p(s)V_s^i(p)$ over E admits a unique non-degenerate sesquilinear form $\langle (u),(v)\rangle$ over E satisfying

(i) $\langle (v),(u)\rangle = (-1)^{i-1}\varepsilon \overline{\langle (u),(v)\rangle}$

(ii) $\mathrm{trace}_{E/F}\langle (u),(v)\rangle = (u, D^{i-1}v)$ for $u,v \in V_s^i(p)$,

where (u) denotes the equivalence class of the vector u in the quotient space. (We assume that char(F) $\neq 2$). The isomorphism class of $(V_s(p),s)$ in Auto(CIPS) determines the equivalence class of the sesquilinear space $H_s^i(p)$ up to multiplication of $\langle\ ,\ \rangle$ by elements of F. (The same element of F must be used for each i). If we are given conformal equivalences $\alpha_i : H_s^i(p) \longrightarrow H_t^i(p)$ for

all i such that $\langle \alpha_i(u), \alpha_i(v) \rangle = k\langle (u),(v) \rangle$ for all appropriate (u), (v) and i with k a fixed non-zero element of F, then there is an isomorphism $w:(V_s(p),s) \longrightarrow (W_t(p),t)$ in Auto(CIPS) with $\lambda(w) = k$.

Proof Since char(F) $\neq 2$, the automorphism of E is not the identity. The existence and uniqueness of $\langle\ ,\ \rangle$ may be proved as in the proof of Theorem 3.3.5. (Since char(F) $\neq 2$, the separability of the field extension E over F is guaranteed. In fact every element of E may be written uniquely as $a+b\xi$ with $a,b \in F$, and this element has trace 2a).

In order to prove the rest of the theorem, we imitate the proof of Corollary 3.3.6, but we do not need to consider the cases $\varepsilon = 1$ and $\varepsilon = -1$ separately. We take an orthogonal basis $(u_1),\ldots,(u_r)$ of $H_s^i(p)$ with $\langle (u_j),(u_j) \rangle = c_j = \bar{c}_j$, a non-zero element of F. Choosing an arbitrary representative u_1 of (u_1), we replace it by $u_1' = u_1 + a(s)Du_1$ where $a(x)$ is a linear polynomial in $F[x]$. Using the methods of Corollary 3.3.6, we eventually show that the representative u_1 could have been chosen so that

$$(s^k D^j u_1, u_1) = 0 \text{ for } k = 0,1 \text{ and } 0 \leqslant j \neq i-1,$$
$$= \operatorname{trace}_{E/F}(\xi^k c_1).(-1)^{i-1} \text{ if } j = i-1.$$

Using the basis with elements $s^k D^j u_1$ for $F[s]u_1$, we do not need to consider polynomials in the variable $x \pm \lambda x^{-1}$, and we may complete the proof in the same way as Corollary 3.4.6.

3.5 Conjugacy classes in the unitary groups

Let F be a field with an automorphism, $\sigma: a \longmapsto \bar{a}$, such that σ^2 = identity. Let $\mathcal{C}$ = HPS be the category whose objects are finite dimensional vector spaces V over F on which is defined a non-degenerate sesquilinear form (,) such that $(v,u) = \overline{(u,v)}$ for all u and v in V. A morphism $s: V \longrightarrow W$ in HPS is an F-linear map such that (su,sv) = (u,v) for all u and v in V. (The same symbol (,) is used for the sesquilinear form on each V. This should not cause confusion). We could discuss the conjugacy classes in the group Aut(V) = U(V), the unitary group of V, in the same way as in 3.3, but this would require yet another repetition of the arguments of 3.3 and 3.4. (A uniform proof along these lines dealing with the orthogonal, symplectic and unitary cases simultaneously is sketched in Chapter IV, section 2 of the article on conjugacy classes by T.A. Springer and R. Steinberg in "Seminar on Algebraic Groups and Related Finite Groups", Lecture Notes in Mathematics 131, Springer Verlag, 1970). The conjugacy classes are found to depend on the equivalence classes of sesquilinear spaces, none of which are F-bilinear. If F is a finite field the equivalence class of such a non-degenerate sesquilinear space is completely determined by its dimension. In this case we may prove the following results which are quoted from G.E. Wall: "On the conjugacy classes in the unitary, symplectic and orthogonal groups", J. Austr. Math. Soc., vol. 3 (1963).

Theorem 3.5.1

Let $F = F_q$, the finite field with $q = r^2$ elements and let $V \in$ HPS. Let $a \longmapsto \bar{a}$ be the automorphism of F given by $\bar{a} = a^r$ for all $a \in F$. Given $g(x) = x^m + a_1 x^{m-1} + \ldots + a_m \in F[x]$ with $g(0) = a_m \neq 0$, define another such polynomial by

$$g^*(x) = \overline{g(0)^{-1}} x^{\deg(g)} \bar{g}(1/x) = (\bar{a}_m x^m + \bar{a}_{m-1} x^{m-1} + \ldots + 1)/\bar{a}_m.$$

Then

(i) Two elements s and t are conjugate in U(V) if and only if they are conjugate in the larger group GL(V).

(ii) An element s of GL(V) is conjugate in GL(V) to some element of the subgroup U(V) if and only if for every monic irreducible p the sequence $(n_1,\ldots,n_k)$ of Theorem 3.2.1(iv) is the same for $V_s(p)$ and $V_s(p^*)$.

For the rest of 3.5 we assume that $F = F_q$ is the finite field with $q = r^2$ elements, that the automorphism of F is as described in Theorem 3.5.1, and that $V \in HPS$ with $\dim_F V = n$.

Proposition 3.5.2

$$|U(V)| = r^N(r+1)(r^2-1)(r^3+1) \ldots (r^n-(-1)^n)$$

where $N = n(n-1)/2$

Proof See (for example) L.E. Dickson: Linear Groups, Leipzig, 1901.

Before stating and proving the result analogous to Theorem 3.2.5 for the unitary groups, we must investigate the possible eigenvalues of elements of U(V).

Lemma 3.5.3

If $g(x)$ is a monic polynomial with $g(0) \neq 0$, and ξ is a root of g, then ξ^{-r} is a root of g^*.

Proof This is immediate from the definition of $g^*(x)$ and from the identity:

$$g(x)^r = \bar{g}(x^r),$$

where the polynomial $\bar{g}$ is obtained from g by applying the field automorphism to its coefficients. (r is a power of the characteristic of F).

Proposition 3.5.4

Let $p(x) = p^*(x)$ be a monic irreducible polynomial in $F[x]$ of degree i. Let ξ be a root of p and ζ be a generator of the (cyclic) multiplicative group of $F[\xi]$, the finite field with q^i elements. Suppose that $\xi = \zeta^a$. Then

(i) deg(p) = i is odd

(ii) a is an integral multiple of r^i-1

(iii) For each odd integer, i, it is possible to choose a monic irreducible polynomial $p(x) = p^*(x)$ of degree i, a root ξ of p, and a generator ζ of the multiplicative group of $F[\xi]$ with $\xi = \zeta^a$ and $a = r^i-1$.

Proof By lemma 3.5.3 $\xi^{-r} = \zeta^{-ar}$ is also a root of p, and so is equal to $\xi^{q^b} = \zeta^{aq^b}$ for some integer b by the Galois theory of finite fields. Since the multiplicative group of $F[\xi]$ is cyclic of order q^i-1 and ζ is a generator, we have the congruence:

$$-ar \equiv aq^b \mod(q^i-1)$$

i.e. $$a(r^{2b}+r) \equiv 0 \mod(q^i-1)$$

Hence $$a(r^{2b-1}+1) \equiv 0 \mod(r^{2i}-1)$$ since $q = r^2$ and $(r, r^{2i}-1)=1$. This means that $a(r^{2b-1}+1)$ is divisible by $r^{2i}-1$, and so a must be divisible by $(r^{2i}-1)/(r^{2i}-1, r^{2b-1}+1) = (r^{2i}-1)/(r^{(2b-1,2i)}+1)$ by Proposition 1.1.3(iv). Thus a must be divisible by $(r^i-1)(r^i+1)/(r^{(2b-1,i)}+1)$. If i is even, the denominator of this fraction divides r^i-1 by Proposition 1.1.3(iv), and so a is a multiple of r^i+1. This implies that $\xi = \zeta^a$ satisfies $\xi^{r^i-1} = \xi^{q^{i/2}-1} = 1$, so that $F[\xi]$ is contained in the field with $q^{i/2}$ elements, contradicting the definition of i. Hence i is odd and we have proved (i).

(ii) now follows immediately since if i is odd, the denominator $r^{(2b-1,i)}+1$ divides r^i+1 by Proposition 1.1.3(iii), and so a must be divisible r^i-1, as required.

Let ζ be a generator of the cyclic multiplicative group of the field with q^i elements with i odd. Let $\xi = \zeta^{r^i-1}$. So ξ has

multiplicative order $(q^i-1)/(r^i-1) = r^i+1$. For $j < i$, we have $(r^i+1, q^j-1) = (r^i+1, r^{2j}-1) < r^i+1$ by Proposition 1.1.3(iv) and inspection, since $r>1$. Thus ξ is not contained in any proper subfield of the field with q^i elements, and so its irreducible polynomial over F has degree i.

To complete the proof of (iii) it is enough to show that $\xi^{-r} = \xi^{q^a}$ for some integer c. So we must show that there is a solution, , of the congruence

$$q^a \equiv -r \bmod(r^i+1),$$

i.e. $$r^{2c}+r \equiv 0 \bmod(r^i+1)$$

i.e. $$r(r^{2c-1}+1) \equiv 0 \bmod(r^i+1).$$

i is odd by hypothesis, and so we may choose c so that $i = 2c-1$. This completes the proof of (iii).

Theorem 3.5.5

Let $F = F_q$ be the finite field with $q = r^2$ elements, and let $V \in HPS$ with $\dim_F V = n$. Let S be the set of integers, k, such that s and s^k are conjugate for all $s \in U(V)$. Then

(i) if $n = 1$, S is the set of integers, k, with $k \equiv 1 \bmod(r+1)$

(ii) if $n>1$, S is the set of integers, k, such that

$$\left.\begin{array}{l} k \equiv q^{a_i} \bmod(r^i+1) \text{ for odd } i \leqslant n \\ k \equiv (-r)^{a_i} \bmod(r^i-1) \text{ for even } i \leqslant n \end{array}\right\} \text{where } a_1, \ldots, a_n \text{ are integers}$$

$$(k,q) = 1.$$

Proof We use the methods of Theorem 3.2.5. Let $s \in U(V)$ and let $p(x)$ be a factor of its characteristic polynomial. By Theorem 3.5.1(ii), $p^*(x)$ is also a factor. By Lemma 3.5.3, we see that if ξ is an eigenvalue of s (in some algebraic extension of F), then so are ξ^{q^a} and ξ^{-rq^a} for all integers $a \geqslant 0$. No other restrictions are placed on the remaining possible eigenvalues by Theorem 3.5.1.

If $p(x)$ is monic irreducible with $p \neq p^*$, we must have $2\deg(p) \leqslant n$ for it to be a possible factor of the characteristic

polynomial of some element of U(V), and so every polynomial of degree less than or equal to $\frac{1}{2}n$ over F may arise as a factor of the characteristic polynomial of some element of U(V). If $2j \leq n$, and ζ is a generator of the multiplicative group of the field with $q^j = r^{2j}$ elements, this shows that ζ is a possible eigenvalue of some $s \in U(V)$. By the remarks in the last paragraph, if $k \in S$, then

$$\zeta^k = \zeta^{q^a} \quad \text{or} \quad \zeta^{-rq^a} \quad \text{for some positive integer a.}$$

Hence $k \equiv q^a = (-r)^{2a} \bmod(q^j-1)$ by the method of Theorem 3.2.5,

or $k \equiv -rq^a = (-r)^{2a+1} \bmod(q^j-1)$.

This gives rise to the second congruence in the statement of the theorem.

If $p = p^*$, Proposition 3.5.4, and the same methods yield the first congruence. The claim that $(k,q) = 1$ follows from Proposition 3.5.2 for $n > 1$ since k must be prime to the order of U(V).

Conversely, if k satisfies the given conditions, it is easy to see that k is prime to the order of every element of U(V) by Proposition 3.5.2, and that for $s \in U(V)$, s and s^k must be conjugate in GL(V) by the argument in the last two paragraphs of the proof of Theorem 3.2.5. Hence they are conjugate in U(V) by Theorem 3.5.1(i). This proves (ii). (i) is a very easy special case.

Corollary 3.5.6

The set S of Theorem 3.5.5(ii) is the set of integers k such that

$$k \equiv (-r)^{a_i} \bmod(r^i-(-1)^i) \text{ for some integers } a_i \text{ and } 1 \leq i \leq n,$$

$$(k,q) = 1$$

Proof The last paragraph of the proof of Proposition 3.5.4 showed that for odd i there is an integer c with $q^c \equiv -r \bmod(r^i+1)$. Replacing $(-r)^{a_i}$ by $q^{c+(a_i-1)/2}$ when both i and a_i are odd shows that the two sets of congruences are equivalent.

Section 4 Calculation of the Fields Generated by the Values of the Characters of the Finite Classical Groups

4.1 The general linear and unitary groups over finite fields

We wish to apply the theory described in Section 2 to the groups whose conjugacy classes have been described in Section 3. We use most of the number theory described in Section 1, and prove more such results as we need them. There are striking similarities between the results for the general linear and the unitary groups defined over finite fields. For this reason we treat the two cases together.

Let $F = F_q$ be the finite field with $q = p^b$ elements where p is a prime number, and let V be a vector space of dimension n over F.

Lemma 4.1.1

The exponent of the group GL(V) is

$$m = p^c[q-1, q^2-1, q^3-1, \ldots, q^n-1]$$

where $c \geq 1$ if and only if $n > 1$. (In fact it is not hard to show that for $n > 1$, $c = 1 + \log_p(n-1)$).

Proof If s^m = identity for all $s \in GL(V)$, then $\xi^m = 1$ for all possible eigenvalues ξ of elements of GL(V). This implies that m is a multiple of q^i-1 for each i between 1 and n, since every irreducible polynomial of degree less than n+1 can be a factor of the characteristic polynomial of some element of GL(V). Hence m is a multiple of their least common multiple: $[q-1, \ldots, q^n-1]$. However any element of GL(V) raised to this power has all its eigenvalues equal to 1 by the same argument. If some $t \in GL(V)$ has all its eigenvalues equal to 1, then t-identity = t-I has all its eigenvalues equal to 0, and so is nilpotent. Using the binomial theorem to expand $t^k = (I+(t-I))^k$, in powers of t-I, we find that for k a sufficiently high power of p, p divides all the binomial coefficients $\binom{k}{i}$ for

which $(t-I)^i \neq 0$. This proves that m has the given form. The assertion about when $c > 1$ follows by inspection, or from Proposition 3.2.4.

Proposition 4.1.2

The set of simultaneous congruences

$$k \equiv q^{a_i} \bmod(q^i-1) \text{ for some integer } a_i, \text{ for } 1 \leqslant i \leqslant n$$

of Proposition 3.2.5 is equivalent to the single congruence

$$k \equiv q^a \bmod(\lceil q-1, q^2-1, \ldots, q^n-1 \rceil) \text{ for some integer a.}$$

Proof By Theorem 1.1.2 the simultaneous congruences have a solution if and only if $q^{a_i} \equiv q^{a_j} \bmod((q^i-1, q^j-1))$ for all i and j,

i.e. $q^{a_i-a_j} \equiv 1 \bmod(q^{(i,j)}-1)$ by Proposition 1.1.3(ii)

i.e. $a_i-a_j \equiv 0 \bmod((i,j))$ by inspection.

By Theorem 1.1.2, again, this happens if and only if there is an integer a with $a \equiv a_i \bmod(i)$ for $1 \leqslant i \leqslant n$. Then $q^a \equiv q^{a_i} \bmod(q^i-1)$ for $1 \leqslant i \leqslant n$ by inspection. The reverse implication is trivial.

Corollary 4.1.3

In the notation of 2.1 with m as in Lemma 4.1.1, $\Gamma(GL(V))$ is the subgroup $G_{ram} \oplus G_{Fr}$ of $(Z/mZ)^*$ where G_{ram} consists of those $k \in (Z/mZ)^*$ such that $k \equiv 1 \bmod(\lceil q-1, \ldots, q^n-1 \rceil)$ and G_{Fr} is the cyclic group generated by the unique $k \in (Z/mZ)^*$ such that

$$k \equiv 1 \bmod(p^c) \text{ and } k \equiv q \bmod(\lceil q-1, \ldots, q^n-1 \rceil).$$

(Compare with Theorem 1.3.5)

Proof This is immediate from Theorem 3.2.5, Lemma 4.1.1 and the preceding proposition.

Corollary 4.1.4

$$Q_q(GL(V)) = Q_q.$$

Hence all the irreducible characters of GL(V) take values in the ring of Witt vectors $W(F_q)$.

Proof Immediate from Corollary 4.1.3 and Theorem 2.2.2.

Note In fact $Q_p(GL(V)) = Q_q$, as can be seen using Theorem 2.2.2 to show that $Q_r(GL(V)) \neq Q_r$ for any proper factor r of q.

Now we examine extensions of the rational numbers.

Theorem 4.1.5

$$Q(GL(V)) = Q(\sqrt[q-1]{1})^q Q(\sqrt[q^2-1]{1})^q \ldots Q(\sqrt[q^n-1]{1})^q$$

where $Q(\sqrt[j]{1})^q$ denotes the subfield of $Q(\sqrt[j]{1})$ fixed pointwise by the automorphism sending w to w^q whenever $w^j = 1$.

Proof Applying Theorem 2.2.1 with m as in Lemma 4.1.1 and $n = p^{-a}m$, to the situation of Corollary 4.1.3, we see that

$Q(GL(V)) = Q(\sqrt[N]{1})^q$ where $N = [q-1, q^2-1, \ldots, q^n-1]$, (where of course n now means $\dim_F V$). (In the notation of Theorem 2.2.1, it is easy to see that $\mathrm{Ker}(f_{m,(m,n)})$ is just the subgroup G_{ram}, and that $f_{m,(m,n)}(\Gamma(G))$ is the cyclic group generated by q in $(Z/NZ)^*$).

We prove by induction on n that $Q(\sqrt[N]{1})^q$ is equal to the compositum of fields in the statement of the theorem. This is trivial for n = 1. Suppose it true for all integers less than n. Let $L = Q(\sqrt[M]{1})$, $K = Q(\sqrt[q^n-1]{1})$ and $M = [q-1, q^2-1, \ldots, q^{n-1}-1]$, so that $L = Q(\sqrt[q-1]{1}, \ldots, \sqrt[q^{n-1}-1]{1})$. Then the order of q restricted to K is n, and the order of q restricted to L is $[1,2,\ldots,n-1]$ by inspection. $L \cap K$ contains the $(q^{(i,n)}-1)^{th}$ roots of unity for $1 \leqslant i \leqslant n-1$ by Proposition 1.2.3 and Proposition 1.1.3(ii). Hence the restriction of q to $L \cap K$ has order at least $[(1,n),(2,n),\ldots,(n-1,n)]$. By Corollary 1.2.6 and Proposition 1.1.1(iv), this shows that $(LK)^q = L^q K^q$. The induction hypothesis shows that L^q has the desired form.

Next we consider the unitary groups. Let $F = F_q$ be the finite field with $q = r^2$ elements, and let $a \mapsto \bar{a} = a^r$ be the automorphism of F. Let V be a finite dimensional vector space over F with a non-degenerate sesquilinear form (,), and write U(V) for the unitary group of V.

Proposition 4.1.6

$$Q_q(U(V)) = Q_q$$

Proof Apply Proposition 2.3.2 to Corollary 4.1.4 and Theorem 3.5.1(i). (Of course this could also be proved directly by the methods already used for GL(V)).

Proposition 4.1.7

The set of simultaneous congruences

$$k \equiv (-r)^{a_i} \bmod(r^i-(-1)^i) \text{ for some integers } a_i \text{ and } 1 \leqslant i \leqslant n$$

of Corollary 3.5.6 is equivalent to the set of simultaneous congruences

$$k \equiv (-r)^a \bmod([r+1, r^2-1, \ldots, r^n-(-1)^n]) \text{ for some integer } a.$$

Proof We first note that for all positive integers i and j

$$(r^i-(-1)^i, r^j-(-1)^j) = r^{(i,j)} - (-1)^{(i,j)}.$$

(This follows from Proposition 1.1.3 by writing "r" for "q" and considering separately the three cases: i and j both odd, i and j both even, and i even with j odd).

For consistency of the congruences of Corollary 3.5.6 we need

$$(-r)^{a_i-a_j} \equiv 1 \bmod(r^{(i,j)}-(-1)^{(i,j)}) \text{ for all } i,j$$

by Theorem 1.1.2. If (i,j) is even = k, say, this gives

$(-r)^{a_i-a_j} \equiv 1 \bmod(r^k-1)$ which is clearly equivalent to

$a_i-a_j \equiv 0 \bmod((i,j))$.

Similarly if (i,j) is odd we find by inspection that the congruence is equivalent to $a_i-a_j \equiv 0 \bmod((i,j))$. So by Theorem 1.1.2 there is an integer a with $a \equiv a_i \bmod(i)$ for $1 \leqslant i \leqslant n$, and it follows that

$$(-r)^a \equiv (-r)^{a_i} \bmod(r^i-(-1)^i) \text{ for } 1 \leqslant i \leqslant n.$$

Proposition 4.1.8

If $q = r^2 = p^{2b}$ where p is a prime number, the exponent of the group $U(V)$ is $m = p^d\lceil r+1, r^2-1, \ldots, r^n-(-1)^n\rceil$ and $d \geqslant 1$ if and only if $n > 1$. With this value of m, $\Gamma(U(V))$ is the subgroup $G_{ram} \oplus G_P$ of $(Z/mZ)^*$ where G_{ram} consists of those $k \equiv 1 \bmod(\lceil r+1, \ldots, r^n-(-1)^n\rceil)$, and G_P is the cyclic group generated by the unique $k \in (Z/mZ)^*$ such that

$$k \equiv 1 \bmod(p^d) \text{ and } k \equiv -r \bmod(\lceil r+1, \ldots, r^n-(-1)^n\rceil).$$

Proof The value of the exponent may be found by the method used in Lemma 4.1.1, using results on the possible eigenvalues of elements of $U(V)$ from the proof of Theorem 3.5.5.

The rest follows from Corollary 3.5.6 and Proposition 4.1.7.

Theorem 4.1.9

$$Q(U(V)) = Q(\sqrt[r+1]{1})^{-r} Q(\sqrt[r^2-1]{1})^{-r} \ldots Q(\sqrt[r^n-(-1)^n]{1})^{-r}$$

where $Q(\sqrt[j]{1})^{-r}$ denotes the subfield of $Q(\sqrt[j]{1})$ fixed pointwise by the automorphism sending w to w^{-r} whenever $w^j = 1$.

Proof The proof is almost identical to that of Theorem 4.1.5 with a few obvious modifications in the notation.

It is interesting to write down generators of the algebraic number fields obtained for $Q(GL(V))$ and $Q(U(V))$ since these might be expected to appear as entries in the character tables of the groups.

For GL(V) we are interested in generators of the fields $Q(\sqrt[q^i-1]{1})^q$. Let ζ be a $(q^i-1)^{th}$ root of unity. Then

$$\zeta + \zeta^q + \zeta^{q^2} + \ldots + \zeta^{q^{i-1}} \in Q(\sqrt[q^i-1]{1})^q,$$

which is spanned as a vector space over Q by such elements as ζ runs through all the $(q^i-1)^{th}$ roots of unity.

For U(V) we are interested in generators of the fields

$Q(\sqrt[i]{r^i-(-1)^i})^{-r}$. Let η be a $(r^i-(-1)^i)^{th}$ root of unity. Then

$$\eta + \eta^{-r} + \eta^{r^2} + \ldots + \eta^{(-r)^{i-1}} \in Q(\sqrt[i]{r^i-(-1)^i}),$$

which is spanned as a vector space over Q by such elements as η runs through all the $(r^i-(-1)^i)^{th}$ roots of unity.

4.2 The orthogonal and symplectic groups - Preliminary results

Here we collect together the solutions of some of the problems which arise in calculating the fields generated by the characters of both the symplectic and the orthogonal groups over finite fields. In order to exploit the general theory of 3.3 in this case we first examine the classification of sesquilinear and bilinear spaces over finite fields.

Proposition 4.2.1

Let $F = F_q$ be the finite field with $q = r^2$ elements and let σ be the field automorphism with $\sigma(a) = \bar{a} = a^r$ for all $a \in F$. If $(V, \langle\ ,\ \rangle)$ is a non-degenerate sesquilinear space over F satisfying

$$\langle v,u\rangle = \overline{\langle u,v\rangle} \quad \text{for all } u,v \in V$$

then V has an orthogonal basis $e_1,\ldots,e_n$ such that

$$\langle e_i,e_i\rangle = 1 \text{ for } 1 \leqslant i \leqslant n.$$

Proof V has an orthogonal basis $f_1,\ldots,f_n$ by Proposition 3.1.6, and $\langle f_i,f_i\rangle = a_i = \bar{a}_i \neq 0$ since the space is non-degenerate. So $a_i \in F_r$ = the fixed subfield of F under σ, and so there exists $b_i \in F$ with $a_i = b_i^{r+1} = b_i\bar{b}_i$ (since the multiplicative groups of F_q and F_r are cyclic of orders r^2-1 and $r-1$ respectively). Define $e_i = b_i^{-1}f_i$.

Corollary 4.2.2

The equivalence class of a non-degenerate sesquilinear space $(V, \langle\ ,\ \rangle)$ over F_q with $q = r^2$ as in the proposition, such that

$$\langle v,u\rangle = \varepsilon\langle u,v\rangle \quad \text{for all } u,v \in V,$$

is uniquely determined by the pair $(\dim_F V, \varepsilon)$. In particular $(V, \langle\ ,\ \rangle)$ and $(V, a\langle\ ,\ \rangle)$ are equivalent for $a \in F_r$, where $a\langle\ ,\ \rangle$ denotes the form $\langle\ ,\ \rangle$ multiplied by the scalar a.

Note ε must satisfy $\varepsilon\bar{\varepsilon} = 1$ by the proof of Proposition 3.1.3.

Proof If $\varepsilon = 1$ this is immediate from Proposition 4.2.1. If $\varepsilon \neq 1$,

$\langle , \rangle$ is conformally equivalent to a form with $\varepsilon = 1$ by Proposition 3.1.3(ii), and the result follows from this case.

Note If F_q is a finite field, it is straightforward to check that the only possible automorphism of order 2 has the form given in the statement of Proposition 4.2.1. Hence the results above classify all possible non-degenerate reflexive sesquilinear spaces over finite fields.

Next we consider bilinear forms over finite fields.

Proposition 4.2.3

If $(V,(\ , \ ))$ is a non-degenerate symplectic space over the finite field F_q, then the equivalence class of $(V,(\ , \ ))$ is completely determined by the (even) integer $\dim_F V$. In particular $(V,(\ , \ ))$ and $(V,a(\ , \ ))$ are equivalent for all $a \in F_q$.

Proof This is just a special case of Proposition 3.1.7.

Proposition 4.2.4

Let $F = F_q$ be the finite field with q elements where q is odd and let V be a finite dimensional vector space over F with a non-degenerate F-bilinear form $(\ , \ )$ such that

$$(u,v) = (v,u) \text{ for all } u,v \in V.$$

Let $e_1,\ldots,e_n$ be an orthogonal basis of V, (which exists by Proposition 3.1.6), with $(e_i,e_i) = a_i$. Let $d(V) = a_1a_2\ldots a_n(F^*)^2$ considered as an element of the group $F^*/(F^*)^2$, a cyclic group of order 2. Then the equivalence class of $(V,(\ , \ ))$ is uniquely determined by the pair $(\dim_F V, d(V))$.

Proof See, for example, Section 62 of O.T. O'Meara: Introduction to Quadratic Forms, Springer-Verlag (1963).

We are now in a position to make the results on conjugacy in 3.3 more precise for F a finite field.

Theorem 4.2.5

Let $F = F_q$ be the finite field with q elements where q is odd. Let $(V,(\ ,\ ))$ be a non-degenerate bilinear space over F such that

$$(v,u) = \varepsilon(u,v) \text{ for all } u,v \in V,$$

where $\varepsilon = 1$ or -1, and let Aut(V) be the group of isometries of $(V,(\ ,\ ))$. Two elements s and t of Aut(V) are conjugate in Aut(V) if and only if

(i) s and t are conjugate in GL(V),

and (ii) for $p(x) = x+1$ or $x-1$, the bilinear spaces $(U_s^i(p),\langle\ ,\ \rangle)$ and $(U_t^i(p),\langle\ ,\ \rangle)$ defined in Theorem 3.3.8 are equivalent (a) for all odd i if $\varepsilon = 1$,

or (b) for all even i if $\varepsilon = -1$.

Proof (i) and (ii) are certainly necessary conditions for s and t to be conjugate. By the results of 3.3 s and t are conjugate in Aut(V) if and only if for each irreducible $p(x) \in F[x]$ we have

(1) if $p \neq p^*$, then $(V_s(\{p,p^*\}),s)$ and $(V_t(\{p,p^*\}),t)$ are isomorphic in Auto(VS) by Theorem 3.3.3(ii) and (iii). This is automatically satisfied if s and t are conjugate in GL(V) by Theorem 3.2.1(iii) and the definition of $V_s(\{p,p^*\})$.

(2) if $p = p^*$ and $\deg(p) \geq 2$, then the sesquilinear spaces $H_s^i(p)$ and $H_t^i(p)$ over $E = F[x]/p(x)F[x]$ defined in Theorem 3.3.5 are equivalent for each i. By Corollary 4.2.2 this just requires that the sesquilinear spaces have the same dimension for each i, which necessarily happens if s and t are conjugate in GL(V).

(3) if $p(x) = x+1$ or $x-1$, then the bilinear spaces $U_s^i(p)$ and $U_t^i(p)$ over F are equivalent for all i. For the symplectic spaces among these we only need the (even) dimensions to be

equal, which necessarily happens if s and t are conjugate in GL(V). By Theorem 3.3.8 we see that the only remaining conditions are (ii)(a) and (ii)(b), which correspond to the cases in which $U_s^i(p)$ is not symplectic.

Note Theorem 4.2.5 is similar to the results in section 2.6 of G.E. Wall's paper: "On the conjugacy classes in the unitary, symplectic and orthogonal groups", J. Austr. Math. Soc., vol. 3 (1963). The results 4.2.1 - 4.2.4 are well-known and can be found, for example, in L.E. Dickson: "Linear Groups", Leipzig, 1901.

Next we examine the possible eigenvalues of elements in the symplectic and orthogonal groups over finite fields. For monic $g(x) \in F[x]$, define $g^*(x)$ as at the beginning of 3.3.

Proposition 4.2.6

Let $F = F_q$ be the finite field with q elements,

(i) If $g(x)$ is monic with $g(0) \neq 0$, and ξ is a root of g, then ξ^{-1} is a root of g^*.

(ii) Let $p(x) = p^*(x)$ be monic irreducible of degree $k \geq 2$. Let ξ be a root of $p(x)$, and ζ a generator of the (cyclic) multiplicative group of $F[\xi]$, the field with q^k elements. Let $\xi = \zeta^a$. Then

(a) $k = 2d$ is even

(b) a is an integral multiple of $q^d - 1$

(c) for each even integer k it is possible to choose a monic irreducible $p(x) = p^*(x)$ of degree k, a root ξ, and a generator ζ of the multiplicative group of $F[\xi]$ such that $\xi = \zeta^a$ and $a = q^d - 1$.

Proof (i) Obvious from the definition of g^*.

(ii)(a) This is Lemma 3.3.4(b).

(b) Since ξ^{-1} is also a root, $\xi^{-1} = \xi^{q^c}$ for some integer c with $0 \leq c < 2d-1$, and so $\xi^{q^c+1} = 1$. Hence

ξ is contained in the field with q^{2c} elements, and so is in the field with $q^{(2c,2d)}$ elements. Thus we must have c = d, and (ii)(b) is an immediate consequence.

(ii)(c) Let ζ be a generator of the multiplicative group of the field with q^{2d} elements, and let $\xi = \zeta^{q^d-1}$, so that ξ has multiplicative order q^d+1. Hence $\xi^{q^d} = \xi^{-1} \neq \xi$, so that the minimal polynomial of ξ over F satisfies $p(x) = p^*(x)$. It is clear by Proposition 1.1.3(iv) that q^d+1 does not divide q^n-1 for any $n < 2d$. Hence deg(p) = 2d as required.

Finally we examine the behaviour of the bilinear spaces in Theorem 3.3.8 when the element s of Aut(V) is replaced by a power s^k of itself.

Proposition 4.2.7

Let V be a finite dimensional vector space over the field F, and let $s \in GL(V)$ have all its eigenvalues equal to 1. Suppose that $(s-I)^i = 0$. Then for all integers k

$$(s^k - s^{-k})^{i-1} = k^{i-1}(s-s^{-1})^{i-1}$$

where k is interpreted as an element of F in the obvious way. The result remains true for all odd integers k if we assume that s has all its eigenvalues equal to -1 and that $(s+I)^i = 0$.

Proof Let $s = I+t$ where $t^i = 0$. Then

$$s^{-1} = I - t + t^2 - \ldots + (-1)^{i-1}t^{i-1} = I - t + t^2 f(t),$$

$$s^k = I + kt + \binom{k}{2}t^2 + \ldots + \binom{k}{i-1}t^{i-1} = I + kt + t^2 g(t),$$

$$s^{-k} = I - kt + \binom{k+1}{2}t^2 - \ldots + \binom{k+i-2}{i-1}t^{i-1} = I - kt + t^2 h(t),$$

where f(x), g(x) and h(x) are polynomials with integer coefficients and we have used the condition: $t^i = 0$. Hence

$$(s-s^{-1})^{i-1} = (2t - t^2 f(t))^{i-1} = (2t)^{i-1} \text{ since } t^i = 0,$$

and $(s^k-s^{-k})^{i-1} = (2kt + t^2(g(t)-h(t)))^{i-1} = (2kt)^{i-1}$ as required.

The case with all eigenvalues -1 follows by writing -s for s.

4.3 The symplectic groups over finite fields of odd characteristic

Let $F = F_q$ be the finite field with q elements where q is odd, and let V be a finite dimensional vector space over F on which there is a non-degenerate bilinear form (,) satisfying

$$(v,u) = -(u,v) \text{ for all } u,v \in V.$$

In particular, $\dim_F V = 2n$ must be even since (,) is non-degenerate. Let Sp(V) be the symplectic group of V, i.e. the group of isometries of $(V,(\ , \ ))$.

Lemma 4.3.1

An element s of GL(V) is conjugate in GL(V) to some element of the subgroup Sp(V) if and only if

(i) for every monic irreducible $p(x) \in F[x]$, the sequence $(n_1,\ldots,n_k)$ of Theorem 3.2.1(iv) is the same for $V_s(p)$ and $V_s(p^*)$ where $p^*(x)$ is defined as in 3.3,

(ii) for $p(x) = x+1$ or $x-1$, each odd integer occurs an even number of times in the sequence $(n_1,\ldots,n_k)$ of Theorem 3.2.1(iv).

Proof See section 2.6 of G.E. Wall: "On the conjugacy classes in the unitary, symplectic and orthogonal groups", J. Austr. Math. Soc., vol. 3 (1963).

Theorem 4.3.2

Let S be the set of integers, k, such that s and s^k are conjugate in Sp(V) for all $s \in Sp(V)$. Then for $\dim_F V = 2n \geqslant 2$, S is the set of integers, k, such that

$$\left.\begin{aligned} k &\equiv \varepsilon_i q^{a_i} \bmod (q^i-1) \\ k &\equiv q^{b_i} \end{aligned}\right\} \text{ for } 1 \leqslant i \leqslant n, \text{ where } a_1,b_1,\ldots,b_n \text{ are integers and each } \varepsilon_i = 1 \text{ or } -1.$$

k = a square non-zero element of F_q.

Proof By Lemma 4.3.1, if p(x) is a monic irreducible factor of the characteristic polynomial of some $s \in Sp(V)$, then so is $p^*(x)$. If

$p(x) \neq p^*(x)$ we must have $2.\deg(p) \leqslant 2n$, and so $i = \deg(p) \leqslant n$. For $k \in S$, s and s^k must have the same eigenvalues, and so if $p(\xi) = 0$, ξ^k must be a root of $p(x)$ or of $p^*(x)$, i.e.

$$\xi^k = \xi^{q^a} \text{ or } \xi^{-q^a} \text{ for some } a.$$

Since $p(x)$ may be any irreducible polynomial of degree $i \leqslant n$, this gives the first set of congruences in the theorem.

If $p(x) = p^*(x)$ we are in the situation of Proposition 4.2.6(ii), and may immediately deduce the second set of congruences in the statement of the theorem. All the congruences are clearly both necessary and sufficient for all elements $s \in Sp(V)$ to have the same eigenvalues as s^k.

Finally we must consider $p(x) = x+1$ or $x-1$. By Theorem 4.2.5 we must consider the equivalence classes of the bilinear spaces $U_s^i(p)$ defined in Theorem 3.3.8, for all even i. By the definition of these spaces and Proposition 4.2.7, replacing s by s^k multiplies the bilinear form by k^{i-1}, and so multiplies $d(U_s^i(p))$ by k^j where $j = (i-1)\dim(U_s^i(p))$. ($d(V)$ was defined for a bilinear space V in Proposition 4.2.4). Since $i-1$ is odd and $\dim(U_s^i(p))$ can be odd by Lemma 4.3.1(ii), (it is just the number of n_c with $n_c = i$), k^j is always a square in F_q only if k is one. The converse is easily proved.

Corollary 4.3.3

The exponent of $Sp(V)$ is

$$m = p^c\left[q-1, q+1, q^2-1, q^2+1, \ldots, q^n-1, q^n+1\right] \text{ where } c > 0.$$

Proof This may be proved by the method used in the proof of Lemma 4.1.1, using the results on possible eigenvalues of elements of $Sp(V)$ from the proof of Theorem 4.3.2.

Corollary 4.3.4

The conditions defining the set S in Theorem 4.3.2 may be rewritten as

$$\left.\begin{aligned} k &\equiv \varepsilon q^{a} \bmod(q^{i}-1) \\ k &\equiv q^{b_i} \bmod(q^{i}+1) \end{aligned}\right\} \text{ for } 1 \leq i \leq n \text{ where } \varepsilon = 1 \text{ or } -1,$$

k = a non-zero square in F_q.

Proof For consistency the first set of congruences in the theorem must satisfy $\varepsilon_i q^{a_i} \equiv \varepsilon_j q^{a_j} \bmod((q^i-1, q^j-1))$ for all i and j by Theorem 1.1.2, and so certainly $\varepsilon_i \equiv \varepsilon_j \bmod(q-1)$ for all i and j. Since $q \geq 3$, this forces ε_i = the same value for all i, and hence

$$q^{a_i} \equiv q^{a_j} \bmod((q^i-1, q^j-1)) \text{ for all i and j.}$$

The proof continues as for Proposition 4.1.2.

Lemma 4.3.5

If p is an odd prime and $a \geq 1$, then the cyclotomic field $Q(\sqrt[p^a]{1})$ contains a unique subfield of degree 2 over Q. The subfield is

$$Q(\sqrt{\varepsilon p}) \text{ where } \varepsilon = (-1)^{(p-1)/2}.$$

Proof $\mathrm{Gal}(Q(\sqrt[p^a]{1})/Q) = (Z/p^aZ)^*$ is cyclic of even order $p^a - p^{a-1}$, and so has a unique subgroup of index 2. By Galois theory this proves the existence of a unique subfield of degree 2 over Q. To show that this subfield has the required form, it is enough to show that $\sqrt{\varepsilon p} \in Q(\sqrt[p]{1})$. We do this by a trick used in some proofs of the Law of Quadratic Reciprocity.

For $x \in (Z/pZ)^*$ define the "Legendre symbol" $(\frac{x}{p})$ to be 1 if x is a square and -1 otherwise, so that the map $x \mapsto (\frac{x}{p})$ is a group homomorphism. Let w be a primitive p^{th} root of unity, and define

$$z = \sum_x (\tfrac{x}{p}) w^x, \text{ the sum being over all } x \in (Z/pZ)^*.$$

Hence $z^2 = \sum_{x,y} (\frac{x}{p})(\frac{y}{p}) w^{x+y}$, the sum being over all $x,y \in (Z/pZ)^*$

$$= \sum_{x,t} (\tfrac{x}{p})(\tfrac{tx}{p}) w^{x(1+t)} \text{ by substituting } y = tx$$

$$= \sum_{x,t} (\tfrac{t}{p}) w^{x(1+t)} = \sum_t \left((\tfrac{t}{p}) \sum_x w^{x(1+t)} \right)$$

Now if $1+t = 0$ in Z/pZ, $\sum_x w^{x(1+t)} = p-1$. If $1+t \neq 0$, then $\sum_x w^{x(1+t)} = \sum_y w^y = -1$ since $1+w+w^2+\ldots+w^{p-1} = 0$. Hence $z^2 = (\frac{-1}{p})(p-1) - \sum_{t\neq -1} (\frac{t}{p}) = (\frac{-1}{p})p$ since $\sum_t (\frac{t}{p}) = 0$. (There are as many non-squares as squares in $(Z/pZ)^*$). The result follows by observing that -1 is a square if and only if 4 divides $p-1$.

Proposition 4.3.6

$$Q_q(Sp(V)) = Q_q(\sqrt{\varepsilon q}) \text{ where } \varepsilon = (-1)^{(q-1)/2}.$$

In particular, if q is a square, then $Q_q(Sp(V)) = Q_q$.

Proof Let m be as in Corollary 4.3.3. By Corollary 4.3.4 $\Gamma(G)$ certainly contains the $k \in (Z/mZ)^*$ such that

$k \equiv 1 \bmod(p^c)$ and $k \equiv q \bmod([q-1, q+1, \ldots, q^n-1, q^n+1])$, and so certainly contains the group G_{Fr} of Theorem 2.2.2(ii). We must consider two cases. Let $G = Sp(V)$.

If q is a square, then every integer k is the square of some element of F_q, and so the final condition on S in Corollary 4.3.4 is equivalent to $(k,q) = 1$. In this case $\Gamma(G)$ also contains the subgroup G_{ram}; and so by Theorem 2.2.2(ii), $Q_q(Sp(V)) = Q_q$.

If q is not a square, then $\Gamma(G)$ meets $G_{Fr} \oplus G_{ram} = Gal(Q_q(\sqrt[m]{1})/Q_q)$ in a subgroup of index 2. (Since $(Z/p^cZ)^*$ is cyclic of even order, precisely half of its elements are squares). Hence $Q_q(G)$ has degree 2 over Q_q. $\sqrt{\varepsilon p} \in Q(G) \subseteq Q_q(G)$ by Lemma 4.3.5, Corollary 4.3.4 and the argument above, where $\varepsilon = (-1)^{(p-1)/2} = (-1)^{(q-1)/2}$ since q is an odd power of p. (This is immediate from the fact that all odd primes are congruent to 1 or 3 modulo 4). The result follows since $Q_q(\sqrt{\varepsilon q}) = Q_q(\sqrt{\varepsilon p}) \neq Q_q$.

Note The field found in the Proposition is a ramified extension of Q_q if q is not a square. There are at least two ways to extend the group $Sp(V)$ in order to remove this ramification. One is to extend

the field F_q to the field F_r where $r = q^2$: then we may extend $(,)$ to $V \otimes F_r$ as an alternating F_r-bilinear form and there is an obvious embedding of $Sp(V)$ into $Sp(V \otimes F_r)$. Also $Q_r(Sp(V \otimes F_r)) = Q_r$ is an unramified extension of Q_p.

Another possible extension is by replacing the group $Sp(V)$ with $CSp(V)$, a situation described in a later section. This is more economical, in the sense that the index of $Sp(V)$ in the larger group is relatively small.

Next we attempt to describe suitable generators of the field $Q(Sp(V))$. Unfortunately the situation here is considerably more complicated than for the general linear and unitary groups.

Corollary 4.3.7

$$Q(Sp(V)) = Q(\sqrt{\varepsilon q})Q(\sqrt[N]{1})^H, \text{ with } \varepsilon = (-1)^{(q-1)/2},$$

where $N = [q-1, q+1, q^2-1, q^2+1, \ldots, q^n-1, q^n+1]$ and H is the subgroup of $(Z/NZ)^*$ consisting of those k such that

$k \equiv \varepsilon' q^a \bmod (q^i-1)$ and $k \equiv q^{b_i} \bmod (q^i+1)$ for some integers $a, b_1, \ldots, b_n$ and $\varepsilon' = \pm 1$, for $1 \leq i \leq n$.

Proof Proposition 4.3.6 shows that $\sqrt{\varepsilon q} \in Q(Sp(V))$. If q is a square, the method used in the first paragraph of the proof of Theorem 4.1.5 gives the result: $Q(Sp(V)) = Q(\sqrt[N]{1})^H$. If, on the other hand, q is not a square, it is clear that $[Q(Sp(V)):Q(\sqrt[N]{1})^H] = 2$, by Galois theory, and so it is enough to show that $\sqrt{\varepsilon q} \notin Q(\sqrt[N]{1})$. However this follows from Lemma 4.3.5 since

$Q(\sqrt[p]{1}) \cap Q(\sqrt[N]{1}) = Q$ by Proposition 1.2.3, where q is a power of the prime p.

In order to prodeed further we try to write the subgroup H of $(Z/NZ)^* = Gal(Q(\sqrt[N]{1})/Q)$ as an internal direct sum of cyclic subgroups. Then we may apply the Galois theory of section 1.2 to the

resulting situation.

Firstly we note that $q^i \equiv -1 \bmod(q^i+1)$ for all i, and so the congruences in Corollary 4.3.7 are equivalent to:

$$k \equiv \varepsilon' q^a \bmod(q^i-1) \text{ and } k \equiv \varepsilon' q^{b_i} \bmod(q^i+1) \text{ for } 1 \leqslant i \leqslant n,$$

where the b_i of the corollary has been replaced by $i+b_i$ if $\varepsilon' = -1$.

Next we use Theorem 1.1.2 to find necessary and sufficient conditions for these simultaneous congruences to be consistent. Let h be the (unique) positive integer with $2^h \leqslant n < 2^{h+1}$, and for $0 \leqslant t \leqslant h$ let E_t be the (non-empty) set of integers i with $1 \leqslant i \leqslant n$ such that 2^t divides i but 2^{t+1} does not. For consistency of the simultaneous congruences above, we require

$$\varepsilon' q^{b_i} \equiv \varepsilon' q^{b_j} \bmod((q^i+1, q^j+1)) \text{ for all i and j,}$$

i.e. $q^{b_i - b_j} \equiv 1 \bmod((q^i+1, q^j+1))$ for all i and j.

Using Proposition 1.1.3(iii) we see that this condition is satisfied automatically unless $v_2(i) = v_2(j)$, i.e. unless $i, j \in E_t$ for some t, in which case the condition becomes

$$q^{b_i - b_j} \equiv 1 \bmod(q^{(i,j)}+1).$$

Inspection shows that this is equivalent to the condition

$$b_i \equiv b_j \bmod(2(i,j)), \text{ i.e. } b_i \equiv b_j \bmod((2i, 2j)).$$

By Theorem 1.1.2, again, this is equivalent to the existence of an integer b such that $b \equiv b_i \bmod(2i)$ for all $i \in E_t$, and hence $q^{b_i} \equiv q^b \bmod(q^i+1)$ for all $i \in E_t$. Hence we may assume for each t that $b_i = b_j = b(t)$ for all i and $j \in E_t$.

For consistency we also need

$$q^a \equiv q^{b_j} \bmod((q^i-1, q^j+1)) \text{ for all i and j.}$$

Using Proposition 1.1.3(iv) and the same techniques as above, we find that this is trivially satisfied unless $v_2(i) > v_2(j)$, and that in this case it is equivalent to $a \equiv b_j \bmod(2(i,j))$, i.e.

$$a \equiv b_j \bmod((i, 2j)), \text{ whenever } v_2(i) > v_2(j).$$

For fixed t this becomes:

$a \equiv b(t) \bmod((i,2j))$ for all $j \in E_t$, and all i divisible by 2^{t+1}. Clearly all the information in these congruences is given by those i between 1 and n which are divisible by 2^{t+1} and no higher power of 2, i.e. by those $i \in E_{t+1}$. So our conditions become

$a \equiv b(t) \bmod((i,2j))$ for all $j \in E_t$ and all $i \in E_{t+1}$.

Hence $a \equiv b(t) \bmod(m(t))$ where $m(t)$ = the least common multiple of the set $(i,2j)$ for $j \in E_t$, $i \in E_{t+1}$, by Theorem 1.1.2. $m(t)$ is certainly a factor of the least common multiple of the set E_{t+1} by inspection. However, if $i \in E_{t+1}$, then $\frac{1}{2}i \in E_t$ by definition, and so $m(t)$ = the least common multiple of the set E_{t+1}.

Summing up, we have shown that the simultaneous congruences of Corollary 4.3.7 are equivalent to the following set:

$k \equiv \varepsilon q^a \bmod(q^i-1)$ for some integer a and $1 \leqslant i \leqslant n$,

$k \equiv \varepsilon q^{b(t)} \bmod(q^j+1)$ for all $j \in E_t$, some integer $b(t)$ and $1 \leqslant t \leqslant h$,

where $b(t) \equiv a \bmod(m(t))$ for each t.

(We assume that $m(h) = 1$, since E_{h+1} is empty and so contributes no conditions in the argument above). It is straightforward to check that these congruences are consistent, using Theorem 1.1.2 and Proposition 1.1.3.

Proposition 4.3.8

$Q(Sp(V)) = Q(\sqrt{\varepsilon q})Q(\sqrt[N]{1})^H$ where $\varepsilon = (-1)^{(q-1)/2}$, $N = [q-1, q+1, \ldots, q^n-1, q^n+1]$, and H is the internal direct sum of the cyclic subgroups of $(Z/NZ)^* = Gal(Q(\sqrt[N]{1})/Q)$ with generators:

I. $k = -1 \in (Z/NZ)^*$

II. $k = q \in (Z/NZ)^*$

III. $k = k_t \in (Z/NZ)^*$ for $0 \leqslant t \leqslant h$, where h is the unique integer such that $2^h \leqslant n < 2^{h+1}$, E_t = the set of integers i with $1 \leqslant i \leqslant n$ which are divisible by 2^t but not by 2^{t+1}, $m(t)$ = the least common multiple of the integers in E_{t+1}, (by convention $m(h)=1$),

and k_t is the unique element of $(Z/NZ)^*$ such that

$k_t \equiv 1 \bmod(q^i-1)$ for $1 \leqslant i \leqslant n$,

$k_t \equiv 1 \bmod(q^i+1)$ for $1 \leqslant i \leqslant n$ with $i \notin E_t$,

$k_t \equiv q^{m(t)} \bmod(q^j+1)$ for $j \in E_t$.

Proof An easy consequence of Corollary 4.3.7, and of the arguments preceding the statement of this Proposition.

Notes 1. Most of the difficulties in the calculation which follows are caused by the generators of type III in this Proposition. The description of $Q(Sp(V))$ would be much simpler if they were absent.

2. The existence of the generator, -1, of type I, shows that $Q(\sqrt[N]{1})^H \subseteq$ the real numbers. Hence we may deduce that every absolutely irreducible character of $Sp(V)$ is real-valued if and only if $q \equiv 1 \bmod(4)$.

3. It is easy to compute the orders of the generators of H given in the Proposition: for example q has order $[2,4,6,\ldots,2n]$, -1 has order 2, and k_t has order $2m(t-1)/m(t)$ where $m(-1)$ is interpreted as the least common multiple of all the odd integers between 1 and n. Hence the degree of $Q(\sqrt[N]{1})^H$ over Q is $\phi(N)$ divided by the product of all of these.

We need several lemmas from field theory before continuing. They are collected together in the next Proposition.

Proposition 4.3.9

(i) If E is a set of positive integers, then

(a) $\left(\prod_{i \in E} Q(\sqrt[q^i-1]{1})\right)^q = \prod_{i \in E} Q(\sqrt[q^i-1]{1})^q$

(b) $\left(\prod_{i \in E} Q(\sqrt[q^i+1]{1})\right)^q = \prod_{i \in E} Q(\sqrt[q^i+1]{1})^q$ provided that all the integers in E are divisible by exactly the same highest power of 2,

(c) $\left(\coprod_{i\in E} Q(\sqrt[(q^{2i}-1)/2]{1})\right)^q = \coprod_{i\in E} Q(\sqrt[(q^{2i}-1)/2]{1})^q$

where $\coprod$ denotes the compositum of fields, and $Q(\sqrt[j]{1})^k$ is the subfield of $Q(\sqrt[j]{1})$ fixed by the automorphism sending w to w^k whenever $w^j = 1$.

(ii) If $E_0,\ldots,E_h$ is a partition of $\{1,2,\ldots,n\}$ define

$K_t = \coprod_{i\in E_t} Q(\sqrt[(q^{2i}-1)/2]{1})$ for $0\leqslant t\leqslant h$. Then

(a) $K^q = (K_0)^q(K_1)^q\ldots(K_h)^q$

(b) $(K^q)^{-1} = ((K_0)^q)^{-1}\ldots((K_h)^q)^{-1}$ provided that $q>3$,

where $K = K_0K_1\ldots K_h$.

Proof (i) We prove (c). (The proofs of (a) and (b) are very similar). The proof is by induction on $|E| = n$, say. The result is trivial for $n = 1$. Suppose the result proved for $|E| < n$, and let m be the largest integer in E. Let K be the field $Q(\sqrt[(q^{2m}-1)/2]{1})$ and L be the compositum of the fields $Q(\sqrt[(q^{2j}-1)/2]{1})$ for $m \neq j\in E$. The order of q acting on K is $2m$ since $q^{2m-1} < (q^{2m}-1)/2$ for $q\geqslant 3$. Similarly the order of q acting on L is the least common multiple of the numbers $2j$ for $m \neq j\in E$, $2k$, say. Clearly $L\cap K$ contains the primitive $(q^{(2m,2j)}-1)/2^{th}$ roots of unity for all $m \neq j\in E$, and so the order of q acting on $L\cap K$ is greater than or equal to the least common multiple of the numbers $(2m,2j)$, i.e. $(2m,2k)$ by Proposition 1.1.1(iv). Apply Corollary 1.2.6 and the induction hypothesis.

(ii)(a) Immediate from (i)(c).

(ii)(b) $(K^q)^{-1} = ((K_0)^q\ldots(K_h)^q)^{-1}$ by (ii)(a). Also $(K_i)^q\cap(K_j)^q \supseteq Q(\sqrt[q-1]{1})$ for all i and j, and so -1 has order 2 on this intersection for $q>3$, since then $q-1>2$. The result follows by repeated applications of Corollary 1.2.6 and Proposition 1.2.4(ii). (Note that all the fields considered are subfields of cyclotomic extensions of Q, and so are automatically normal separable extensions

(c) $\left(\coprod_{i\in E} Q(\sqrt[(q^{2i}-1)/2]{1})\right)^q = \coprod_{i\in E} Q(\sqrt[(q^{2i}-1)/2]{1})^q$

where $\coprod$ denotes the compositum of fields, and $Q(\sqrt[j]{1})^k$ is the subfield of $Q(\sqrt[j]{1})$ fixed by the automorphism sending w to w^k whenever $w^j = 1$.

(ii) If $E_0,\ldots,E_h$ is a partition of $\{1,2,\ldots,n\}$ define

$K_t = \coprod_{i\in E_t} Q(\sqrt[(q^{2i}-1)/2]{1})$ for $0\leqslant t\leqslant h$. Then

(a) $K^q = (K_0)^q(K_1)^q\ldots(K_h)^q$

(b) $(K^q)^{-1} = ((K_0)^q)^{-1}\ldots((K_h)^q)^{-1}$ provided that $q>3$,

where $K = K_0K_1\ldots K_h$.

Proof (i) We prove (c). (The proofs of (a) and (b) are very similar). The proof is by induction on $|E| = n$, say. The result is trivial for $n = 1$. Suppose the result proved for $|E|<n$, and let m be the largest integer in E. Let K be the field $Q(\sqrt[(q^{2m}-1)/2]{1})$ and L be the compositum of the fields $Q(\sqrt[(q^{2j}-1)/2]{1})$ for $m\neq j\in E$. The order of q acting on K is 2m since $q^{2m-1} < (q^{2m}-1)/2$ for $q\geqslant 3$. Similarly the order of q acting on L is the least common multiple of the numbers 2j for $m\neq j\in E$, 2k, say. Clearly $L\cap K$ contains the primitive $(q^{(2m,2j)}-1)/2^{th}$ roots of unity for all $m\neq j\in E$, and so the order of q acting on $L\cap K$ is greater than or equal to the least common multiple of the numbers (2m,2j), i.e. (2m,2k) by Proposition 1.1.1(iv). Apply Corollary 1.2.6 and the induction hypothesis.

(ii)(a) Immediate from (i)(c).

(ii)(b) $(K^q)^{-1} = ((K_0)^q\ldots(K_h)^q)^{-1}$ by (ii)(a). Also $(K_i)^q\cap(K_j)^q\supseteq Q(\sqrt[q-1]{1})$ for all i and j, and so -1 has order 2 on this intersection for $q>3$, since then $q-1>2$. The result follows by repeated applications of Corollary 1.2.6 and Proposition 1.2.4(ii). (Note that all the fields considered are subfields of cyclotomic extensions of Q, and so are automatically normal separable extensions

of Q).

For the rest of 4.3 we assume that $q > 3$.

Applying Proposition 4.3.9 to Proposition 4.3.8, we see that

$$Q(Sp(V)) = (((K_0)^q)^{-1}\ldots((K_h)^q)^{-1})^{\langle k_1,\ldots,k_h\rangle}$$

where $\langle k_1,\ldots,k_h\rangle$ is the subgroup of $(Z/NZ)^*$ generated by $k_1,\ldots,k_h$ and K_i is as in Proposition 4.3.9(ii) with E_i as in Proposition 4.3.8. (It is easy to see that $Q(\sqrt[q^i-1]{1})Q(\sqrt[q^i+1]{1}) = Q(\sqrt[(q^{2i}-1)/2]{1})$). Repeated applications of Proposition 1.2.4(ii) show that this field is just $M_0M_1\ldots M_h$ where $M_t = ((K_t^q)^{-1})^{k_t}$. We shall eventually simplify this description.

$K_t = L_tN_t$ with $L_t \cap N_t = Q$ by Proposition 1.2.3, where L_t is the compositum of all the fields $Q(\sqrt[q^i-1]{1})$ for $i \in E_t$ and N_t of all $Q(\sqrt[q^i+1]{1})$ for all $i \in E_t$. Hence it is easily seen that

$$M_t = ((L_t(N_t)^{k_t})^q)^{-1} \text{ since } k_t \text{ acts trivially on } L_t.$$

Lemma 4.3.10

(i) $\left[(L_t(N_t)^{k_t})^q : ((L_t)^q(N_t)^q\right] = \frac{1}{2}m(t)$ if $t < h$

$= 1$ if $t = h$.

(ii) For $k \in E_t$ and w a primitive $(q^{2k}-1)/2^{th}$ root of unity,

(a) the degree of w over $(L_t)^q(N_t)^q$ is $2k^2$

(b) the degree of w over $(L_t(N_t)^{k_t})^q$ is $4k^2/(2k,m(t))$ for $t < h$.

Proof (i) The order of q acting on $L_t \cap (N_t)^{k_t} = Q$ is 1, and the order of q acting on L_t is the least common multiple of the integers in E_t. By definition k_t acts on N_t as $q^{m(t)}$, and the order of q on N_t is the least common multiple of the numbers 2j for $j \in E_t$, which is divisible by m(t). Hence the order of q on $(N_t)^{k_t}$ is m(t), and so

(order of q on L_t, order of q on $N_t^{k_t}$) = (L.C.M. of E_t, m(t))

$= \frac{1}{2}m(t)$ if $t < h$,

$= 1$ otherwise.

By Proposition 1.2.5 this is equal to $\left[(L_t(N_t)^{k_t})^q:((L_t)^q((N_t)^{k_t})^q\right]$ and (i) follows on noting that $((N_t)^{k_t})^q = (N_t)^q$.

(ii)(a) Let $L = (L_tN_t)^H$ where H is generated by σ and τ corresponding to the congruences σ : $k' \equiv 1 \bmod(q^j-1)$, $k' \equiv q \bmod(q^j+1)$ for $j \in E_t$

τ: $k' \equiv q \bmod(q^j-1)$, $k' \equiv 1 \bmod(q^j+1)$ for $j \in E_t$.

(These congruences are consistent by Theorem 1.1.2 and Proposition 1.1.3). Proposition 1.2.4(ii) shows that $L = (L_t)^q(N_t)^q$. Let $K = Q(w) \subseteq L_tN_t$. Then $L \cap K = K^H$ by Proposition 1.2.4(i), and $LK = L(w)$. Hence

$$\begin{aligned}\left[L(w):L\right] &= \left[LK:L\right] = \left[K:L \cap K\right] \text{ by Lemma 1.2.2}\\ &= \left[K:K^H\right] = \text{the order of H acting on K}\\ &= (\text{order of } \sigma \text{ on K}).(\text{order of } \tau \text{ on K}) = 2k.k = 2k^2,\end{aligned}$$

since the cyclic subgroups generated by σ and τ do not intersect, (except in the identity).

(ii)(b) Let $(L_tN_t)^H$ where H is generated by σ and τ, but this time σ : $k' \equiv q \bmod(q^j \pm 1)$ for all $j \in E_t$

τ : $k' = k_t$.

The proof resembles that of (ii)(a), and we need the order of H acting on $K = Q(w)$. The order of σ on K is 2k, and the order of τ is $2k/(2k,m(t))$. The only possible intersection of the two cyclic subgroups (other than the identity) must involve $\sigma^k = q^k$ which is of order 2 on K. ($q^k \equiv 1 \bmod(q^k-1)$). However, for $t \neq h$ the order of τ on K is odd since exactly the same highest power of 2 divides 2k and m(t). Hence in this case the two cyclic subgroups do not intersect.

Note If $t = h$ we just need to change the degree in Lemma 4.3.10(b) to $2k^2/(2k,m(t))$ since it is easy to see that in this case the cyclic subgroups intersect in a group of order 2. (Of course $m(h) = 1$, and so the degree becomes $2k^2$).

Proposition 4.3.11

$$(L_t(N_t)^{k_t})^q = \prod_{k\in E_t} (Q((q^{2k}-1)/2\sqrt{1})^{k_t})^q$$

where L_t and N_t are the fields defined before Lemma 4.3.10.

Proof The field on the left certainly contains that on the right. Let $K = (L_t)^q(N_t)^q$, and let H be the subgroup of $Gal(L_tN_t/K)$ generated by q and k_t. Let $k\in E_t$ and let $w = w_k$ be a primitive $(q^{2k}-1)/2^{th}$ root of unity. Then

$$\begin{aligned}[K(w)^H:K] &= [K(w):K] / [K(w):K(w)^H] \\ &= 2k^2/\text{order of H on K(w)} \text{ by Lemma 4.3.10(ii)(a)} \\ &= 2k^2/(4k^2/(2k,m(t)) \text{ by the proof of Lemma 4.3.10(ii)(b)} \\ &= (2k,m(t))/2. \text{ (If } t = h \text{ we get 1 here instead).}\end{aligned}$$

Now $K(w)^H \subseteq (L_tN_t)^H = (L_t(N_t)^{k_t})^q$, but letting k run over E_t we see

$$\left[\prod_{k\in E_t} K(w_k)^H : K\right] \geq \text{L.C.M. of the } (2k,m(t))/2 \text{ for } k\in E_t$$

But the right-hand side is just $[(L_t(N_t)^{k_t})^q:K]$ by inspection and Lemma 4.3.10(i). (A slight modification shows that this still works for $t = h$). Hence

$(L_t(N_t)^{k_t})^q$ = the compositum of all $K(w_k)^H$ for $k\in E_t$.

However $K = K^H$, and so

$K(w)^H = (K.Q(w))^H = K.Q(w)^H$, and so

$(L_t(N_t)^{k_t})^q = K.$(the compositum of all $Q(w)^H$ for $k\in E_t$)

However applying Proposition 4.3.9(i)(a) and (b) to $(L_t)^q$ and $(N_t)^q$ respectively we find that their compositum K is contained in the compositum of all the $Q(w_k)^H$, and this compositum is just the result we want.

Corollary 4.3.12

$$M_t = \left(\prod_{k\in E_t} (Q((q^{2k}-1)/2\sqrt{1})^{k_t})^q\right)^{-1}$$

$$= \prod_{k \in E_t} ((Q(\sqrt[(q^{2k}-1)/2]{1})^{k_t})^q)^{-1}$$

Proof Only the last line needs justification. The method used to prove Proposition 4.3.9(ii)(b) works here. (Again we use the fact that $q > 3$).

Corollary 4.3.13

$$Q(Sp(V)) = Q(\sqrt[(q^2-1)/2]{1})^{H_1} \ldots Q(\sqrt[(q^{2n}-1)/2]{1})^{H_n}(\sqrt{\varepsilon q}), \quad \varepsilon = (-1)^{\frac{1}{2}(q-1)},$$

where H_k is the subgroup of the Galois group

$$Gal(Q(\sqrt[(q^{2k}-1)/2]{1})/Q) = (Z/\tfrac{1}{2}(q^{2k}-1)Z)^*$$

generated by the three elements q, -1, and s_k of $(Z/\frac{1}{2}(q^{2k}-1)Z)^*$ such that $s_k \equiv 1 \bmod(q^j-1)$ and $s_k \equiv q^{a_k} \bmod(q^j+1)$ for all $j \in E(k)$, where $E(k)$ = the set of integers between 1 and n divisible by exactly the same highest power of 2 as k, and $a_k = (2k,m)$ where m is the least common multiple of the integers between 1 and n which are divisible by a higher power of 2 than k. ($m=1$ if there are no such integers)

Proof This is just a restatement of Proposition 4.3.8, using the later results.

This seems to be the neatest form resembling the results of Theorems 4.1.5 and 4.1.9 in which the field $Q(Sp(V))$ can be expressed. Generators of the fields

$$Q(\sqrt[(q^{2k}-1)/2]{1})^{H_k}$$

appearing in the last Corollary may be written down quite easily, (using the fact that the trace map is surjective), but they do not seem to take on a very simple form.

4.4 The orthogonal groups over finite fields of odd characteristic

Let $F = F_q$ be the finite field with q elements where q is odd, and let V be a finite dimensional vector space over F on which there is a non-degenerate bilinear form $(\ ,\ )$ satisfying

$$(v,u) = (u,v) \text{ for all } u,v \in V.$$

Let O(V) be the orthogonal group of V, i.e. the group of isometries of $(V,(\ ,\ ))$. There are essentially three cases to consider:

I. $\dim_F V$ is odd. Although there are two equivalence classes of non-degenerate bilinear form on V, there is only one conformal equivalence class, and so the orthogonal group O(V) is determined up to isomorphism by $\dim_F V$. (All this may be proved using Proposition 4.2.4).

II. and III. If $\dim_F V$ is even, there are two conformal equivalence classes of non-degenerate forms on V (depending on whether d(V) is a square or not), and in general they give rise to non-isomorphic orthogonal groups. If $\dim_F V = 2n$, we denote these groups by $O_+(V)$ and $O_-(V)$ where $O_+(V)$ arises if V has a basis $e_1,\ldots,e_n,f_1,\ldots,f_n$ such that $(e_i,e_j) = 0 = (f_i,f_j)$ for all i and j, and $(e_i,f_j) = 1$ when $i=j$ and 0 otherwise. $O_-(V)$ corresponds to the other equivalence class of bilinear forms.

Lemma 4.4.1

An element s of GL(V) is conjugate in GL(V) to an element of some orthogonal group O(V) if and only if

(i) for every monic irreducible $p(x) \in F[x]$, the sequence $(n_1,\ldots,n_k)$ of Theorem 3.2.1(iv) is the same for $V_s(p)$ and $V_s(p^*)$ where $p^*(x)$ is defined as in 3.3,

(ii) for $p(x) = x+1$ or $x-1$, each even integer occurs an even number of times in the sequence $(n_1,\ldots,n_k)$ of Theorem 3.2.1(iv).

Suppose next that $n = \dim_F V$ is even and that (i) and (ii) are satisfied. Given $p(x)$, i and $s \in GL(V)$ write $m(p^i)$ for the number of n_j equal to i, in the notation of Theorem 3.2.1(iv). If $m(p^{2i+1}) > 0$ for some i with $p(x) = x+1$ or $x-1$, then s is conjugate in $GL(V)$ to both an element of $O_+(V)$ and an element of $O_-(V)$. If $m((x\pm1)^{2i+1}) = 0$ for all i, then s is conjugate to an element of $O_+(V)$ if and only if $\sum_{i,p} i.m(p^i)$ is even, and s is conjugate in $GL(V)$ to an element of $O_-(V)$ if and only if the same sum is odd.

Proof See section 2.6 of G.E. Wall: "On the conjugacy classes in the unitary, symplectic and orthogonal groups", J. Austr. Math. Soc., vol. 3 (1963).

Theorem 4.4.2

Let S be the set of integers, k, such that s and s^k are conjugate in $O(V)$ for all $s \in O(V)$. Assume that $\dim_F V > 1$.

I. If $\dim_F V = 2n+1$ is odd, S is the set of integers, k, such that

$$\left.\begin{aligned} k &\equiv \varepsilon_i q^{a_i} \bmod (q^i-1) \\ k &= q^{b_i} \bmod (q^i+1) \end{aligned}\right\} \text{ for } 1 \leq i \leq n, \text{ where the } a_i \text{ and } b_j \text{ are integers and each } \varepsilon_i = \pm 1.$$

$(k,q) = 1$.

II. If $\dim_F V = 2n$ and $O(V) = O_+(V)$, the list is the same as in I, except that the congruence modulo (q^n+1) is omitted.

III. If $\dim_F V = 2n$ and $O(V) = O_-(V)$, we omit the congruence modulo (q^n-1).

Proof Similar to that of Theorem 4.3.2. (If $\dim_F V = 1$, $O(V)$ is cyclic of order 2).

The methods of proof of 4.3 yield many results for the orthogonal groups analogous to those already proved for the symplectic groups. We state the most important without details of their proofs.

Theorem 4.4.3

For all three types of orthogonal group over F_q

$$Q_q(O(V)) = Q_q.$$

Proof Much easier than Proposition 4.3.6.

Note We describe the situation in the conformal orthogonal groups in the next section.

Proposition 4.4.4

If $\dim_F V = 2n+1 > 1$, then $Q(O(V))$ is the same as the field described in Corollary 4.3.13 for the symplectic groups, with the term $\sqrt{\varepsilon q}$ omitted.

Proof Compare Theorem 4.4.2, part I, with Theorem 4.3.2.

Proposition 4.4.5

If $\dim_F V = 2n$, then $Q(O(V)) = Q(\sqrt[N]{1})^H$ where H is the internal direct sum of the cyclic subgroups of $(Z/NZ)^* = \mathrm{Gal}(Q(\sqrt[N]{1})/Q)$ with generators $-1, q$ and k_t defined in a way similar to that of Proposition 4.3.8, where

(i) If $O(V) = O_+(V)$, $N = [q-1, q+1, \ldots, q^{n-1}-1, q^{n-1}+1, q^n-1]$,

(ii) If $O(V) = O_-(V)$, $N = [q-1, q+1, \ldots, q^{n-1}-1, q^{n-1}+1, q^n+1]$.

(The sets E_t and the numbers $m(t)$ are modified in ways that become obvious during the proof)

Proof Similar to Proposition 4.3.8.

The calculations can be continued beyond this point, but nothing essentially new emerges.

4.5 The conformal symplectic and orthogonal groups over finite fields of odd characteristic

Let $F = F_q$ be the finite field with q elements where q is odd, and let V be a finite dimensional vector space over F on which there is a non-degenerate bilinear form $(\ ,\ )$ satisfying

$$(v,u) = \varepsilon(u,v) \text{ for all } u,v \in V, \text{ where } \varepsilon = 1 \text{ or } -1.$$

Let $CO(V)$ be the conformal isometry group of $(V,(\ ,\ ))$ if $\varepsilon = 1$, and $CSp(V)$ be the group if $\varepsilon = -1$. Our main interest is in the fields $Q_q(G)$ for these two groups, and so we shall not need the detailed type of information for these groups which we used for the others.

Proposition 4.5.1

Let V be a finite dimensional vector space over the field F, and let $s \in GL(V)$ have all its eigenvalues equal to a. Suppose that $(s-aI)^i = 0$. Then for all integers k such that $a^k = a$,

$$(s^k - a^2 s^{-k})^{i-1} = k^{i-1}(s - a^2 s^{-1})^{i-1}$$

where k is interpreted as an element of F in the obvious way.

Proof Let $s = at$ and apply Proposition 4.2.7 to t.

For the conformal groups this result plays the part that Proposition 4.2.7 played for the symplectic and orthogonal groups. The result analogous to Theorem 4.2.5 is:

Theorem 4.5.2

Two elements s and t of the conformal isometry group of $(V,(\ ,\ ))$ are conjugate in this group if and only if

(i) s and t are conjugate in GL(V),

and (ii) for $p(x) = x-a$ with $a^2 = \lambda(s) = \lambda(t)$, the bilinear spaces defined in Theorem 3.4.7, $U_s^i(p)$ and $U_t^i(p)$, are

conformally equivalent in the way described at the end of Theorem 3.4.7 (a) for all odd i if $\varepsilon = 1$ or (b) for all even i if $\varepsilon = -1$.

Proof Similar to Theorem 4.2.5.

Corollary 4.5.3

Let $s \in \mathrm{Aut}(V)$ = the conformal isometry group of $(V,(\ , \ ))$, and let k be an integer with $k \equiv 1 \bmod(q-1)$. Then s and s^k are conjugate in Aut(V) if and only if they are conjugate in GL(V).

Proof Certainly $a^k = a$ for all $a \in F_q$, and so we may apply Proposition 4.5.1 to the definition of the bilinear spaces in Theorem 3.4.7. The only possible source of difficulty is in the case $\varepsilon = -1$, k not a square in F_q, but in this case k^{i-1} is not a square for all even i and Theorem 4.5.2 is satisfied with $t = s^k$.

Theorem 4.5.4

$$Q_q(\mathrm{CO}(V)) = Q_q$$
$$Q_q(\mathrm{CSp}(V)) = Q_q.$$

Proof In the notation of Corollary 4.5.3, let m be any common multiple of the exponents of Aut(V) and GL(V). In the notation of 2.1, $\Gamma(\mathrm{Aut}(V)) \supseteq \Gamma(\mathrm{GL}(V))$ by Corollary 4.5.3, since every $k \in \Gamma(\mathrm{GL}(V))$ satisfies $k \equiv 1 \bmod(q-1)$ by Corollary 4.1.3. In particular $K(\mathrm{Aut}(V)) \subseteq K(\mathrm{GL}(V))$ for all fields K (of characteristic zero) by Theorem 2.1.1. The result follows by Corollary 4.1.4.

Note The treatment above illustrates how the proofs for other families of groups could have been shortened if we were concerned only with the fields $Q_q(G)$.

4.6 The special orthogonal groups over finite fields of odd characteristic

Let $(V,(\ ,\ ))$ be a non-degenerate symmetric bilinear space over the field F and let O(V) be its orthogonal group. Given $s \in O(V)$ we must have $\det(s) = 1$ or -1, (since if $\dim_F V = n$, $(\ ,\ )$ induces a non-degenerate symmetric bilinear form on the one-dimensional space $\lambda^n V$, and s induces an isometry of this which is just scalar multiplication by $\det(s)$). It is easy, (using the existence of an orthogonal basis if $1 \neq -1$), to show that both 1 and -1 occur as values of $\det(s)$ for $s \in O(V)$. We define the special orthogonal group, SO(V), to be the kernel of the group homomorphism

$$\det: O(V) \longrightarrow \{1,-1\} \subseteq F^*.$$

In particular SO(V) is a normal subgroup of O(V) of index 2, provided that the characteristic of F is not 2.

Note This situation does not arise for the symplectic groups, Sp(V), since it may be shown that $\det(s) = 1$ for all $s \in Sp(V)$.

Now let $F = F_q$ be the finite field with q elements where q is odd. Applying Theorem 2.3.3 with $p = 2$, $G = O(V)$, $H = SO(V)$ and $K = Q_q$, we have

$$\left[\overline{Q_q}(SO(V)) : \overline{Q_q}\right] = 2^k \quad \text{for some integer } k \geqslant 0,$$

by Theorem 4.4.3. However we may use the methods of Theorem 2.3.3 together with our knowledge of conjugacy classes in O(V) to deduce more precise results. We assume that $\dim_F V > 1$.

Theorem 4.6.1

If $\dim_F V$ is odd, then $Q(SO(V)) = Q(O(V))$, as given in Proposition 4.4.4, and $Q_q(SO(V)) = Q_q$.

Proof Consider the normal subgroup, $\{I,-I\}$, of O(V). Since dim(V) is odd, $\det(-I) = (-1)^{\dim(V)} = -1$, and so $SO(V) \cap \{I,-I\} = \{I\}$. Hence O(V) is isomorphic to the Cartesian product group of these two normal

subgroups. The result now follows by Theorem 4.4.3, Proposition 4.4.4 and Theorem 2.3.1(i).

Unfortunately the proof above breaks down if V has even dimension over F.

Theorem 4.6.2

Let $\dim_F V = 2n$ be even, $s \in O(V)$, and m = the exponent of O(V), so that m is certainly a multiple of the exponent of SO(V).

(i) det(s) = -1 if and only if -1 has odd multiplicity as an eigenvalue of s.

(ii) $\Gamma(SO(V)) \subseteq \Gamma(O(V))$ and $[\Gamma(O(V)):\Gamma(SO(V))] = 2^k$ for some integer k.

(iii) $\Gamma(SO(V))$ contains those $k \in (Z/mZ)^*$ such that

$$\left.\begin{array}{l} k \equiv q^{a_i} \bmod(q^i-1) \\ k \equiv q^{2b_j} \bmod(q^j+1) \\ k \text{ is a non-zero square mod}(p) \end{array}\right\}$$

where a_i and b_j are integers and the ranges of the integers i and j are as in Theorem 4.4.2, parts II and III.

Note It is not hard to see from the results in 4.4 that $m = p^c N$, where p is the characteristic of $F = F_q$, N is the integer described in Proposition 4.4.5, and $c > 0$ for $n \geq 1$. Once (i) is established it is clear that m is also the exponent of SO(V).

Proof (i) det(s) = the product of the eigenvalues of s. Let p(x) be a monic irreducible polynomial over F dividing the characteristic polynomial of s, and define $p^*(x)$ as in 3.3. If $p(x) \neq p^*(x)$, then the product of all their roots is 1 by the definition of $p^*(x)$. If $p(x) = p^*(x)$ and $\deg(p) \geq 2$, then the product of the roots of p(x) is 1 by Lemma 3.3.4(b) since deg(p) is even. This leaves the eigenvalues 1 and -1, and the result follows by Lemma 4.4.1.

(ii) The first statement is immediate from (i) and the type of calculation needed to prove Theorem 4.4.2. The second statement

follows from the first and Theorem 2.3.3(i) with p = 2.

(iii) First note that if p(x) is irreducible of degree i and $k \equiv q^a \bmod(q^i-1)$ for some integer a, with $k \in \Gamma(O(V))$, then $V_s(p) = V_t(p)$ where $t = s^k$. (This holds since s and s^k are conjugate in $O(V) \subseteq GL(V)$ and must have the same eigenvalues on the s-invariant subspace $V_s(p)$ by the Galois theory of finite fields). A similar remark applies if $p(x) = p^*(x)$ is irreducible of degree 2j and $k \equiv q^b \bmod(q^j+1)$.

Next note that the values of k in the statement of (iii) are in $\Gamma(O(V))$ by Theorem 4.4.2 and necessarily satisfy the other conditions of the preceding paragraph. Assume k is as in the statement of (iii), and let $s \in SO(V)$. In order to prove that s and $t = s^k$ are conjugate in SO(V), it is sufficient by the remarks above to show that their restrictions are conjugate in the special orthogonal groups of the non-degenerate subspaces: $V_s(p)$ for monic irreducible polynomials with $p(x) = p^*(x)$, and $V_s(\{p,p^*\})$ for those with $p(x) \neq p^*(x)$. (This is clear by Theorem 3.3.3(i) and (ii) and standard properties of the determinant).

Examining closely the proof of Theorem 3.3.3(iii), we see that the isometry k of $V_s(\{p,p^*\})$ constructed there has matrix of the form $\begin{pmatrix} T & 0 \\ 0 & (T')^{-1} \end{pmatrix}$ with respect to a basis of $V_s(p)$ and the dual basis of $V_s(p^*)$ induced by the inner product (,). Hence the isometry has determinant 1 and the restrictions of s and t are automatically conjugate in $SO(V_s(\{p,p^*\}))$.

Next we let p(x) = x-1 and examine the proof of Theorem 3.3.8. (Note that we are interested in the case where i is odd). Choose orthogonal bases $(u_1),\ldots,(u_r)$ of $U_s^i(p)$ and $(v_1),\ldots,(v_r)$ of $U_t^i(p)$, such that the map $(u_j) \longmapsto (v_j)$ induces an equivalence of the two sesquilinear spaces. This induces an isomorphism

$$w:(V_s(p),s) \longrightarrow (V_t(p),t) = (V_s(p),s^k)$$

in Auto(IPS). If $\det(w) = 1$, there is nothing to prove. If $\det(w) = -1$, replace the orthogonal basis of $U_s^i(p)$ by the new basis $(-u_1),(u_2),\ldots,(u_r)$ and follow through the same construction. Since $F[s]u_1$ has dimension i which is odd, this has the effect of changing the sign of $\det(w)$. Hence s and s^k are necessarily conjugate in $SO(V_s(p))$. A similar proof works for $p(x) = x+1$.

Unfortunately the proof of Corollary 3.3.6 cannot be modified in the same way. Let $p(x) = p^*(x)$ be irreducible of degree 2j. Then the restrictions of s and s^k are conjugate in $O(V_s(p))$ for any integer k such that $(k,q) = 1$ and $k \equiv q^a \bmod(q^j+1)$ for some integer a. The argument in the second paragraph of the proof of Theorem 2.3.3 shows that the restrictions of s and s^{k^2} are conjugate in $SO(V_s(p))$. Hence the restrictions of s and s^k are conjugate in $SO(V_s(p))$ provided that $(k,q) = 1$, $k \equiv$ a square mod(p) (which is equivalent to k being a square in $(Z/p^cZ)^*$), and $k = q^{2b} \bmod(q^j+1)$ for some integer b. Clearly the integers described in the statement of (iii) satisfy this.

Corollary 4.6.3

$$Q_q(SO(V)) \subseteq Q_r(\sqrt{\varepsilon p})$$

where $r = q^2$ and $\varepsilon = (-1)^{(p-1)/2}$ with p = characteristic of F_q.

Proof By Theprem 4.6.2(iii) $\Gamma'(SO(V))$ contains those $k \in (Z/p^cN)^*$ such that $k \equiv q^2 \bmod(N)$, $(k,q) = 1$, and k is a square in the cyclic group $(Z/p^cZ)^*$. The result follows from Lemma 4.3.5 in a manner similar to the proof of Proposition 4.3.6.

Notes 1. $[Q_r(\sqrt{\varepsilon p}):Q_q] = 4$, and so the only possibilities for the field $Q_q(SO(V))$ are Q_q, Q_r, $Q_q(\sqrt{\varepsilon p})$ and $Q_r(\sqrt{\varepsilon p})$, and $Q_q(\zeta\sqrt{\varepsilon p})$ where ζ is a primitive $(q^r-1)^{st}$ root of unity.

2. The methods used in the proof of Theorem 4.6.2 could be extended to derive more information about the conjugacy classes in SO(V).

Section 5 Representations of the Weyl Group $W(C_n)$

5.1 The rings R(C) and R(S)

The Weyl group of type C_n, $W(C_n)$ may be defined for $n \geq 0$ as follows:

$W(C_n)$ = the set of permutations t of the set $\{\pm 1, \pm 2, \ldots, \pm n\}$ such that $t(-i) = -t(i)$ for $1 \leq i \leq n$.

We define $W(C_0) = \{I\}$, the group with one element (which may be considered as a group of permutations of the empty set).

For all i and $j \geq 0$ we have an embedding of groups

$$W(C_i) \times W(C_j) \longrightarrow W(C_{i+j})$$

defined by letting $W(C_i)$ act on the set $\{\pm 1, \ldots, \pm i\}$ and $W(C_j)$ act on $\{\pm(i+1), \ldots, \pm(i+j)\}$ in the obvious way.

For each integer $n \geq 0$ let $R(C_n)$ be the representation ring of the group $W(C_n)$, i.e. the Grothendieck group of the category of all complex $W(C_n)$-modules with the product induced by tensor products over the complex numbers. We sometimes identify $R(C_n)$ with the character ring of $W(C_n)$, and identify $R(C_0)$ with the ring Z of ordinary integers in the obvious way. For i and $j \geq 0$ we define the "outer product"

$$R(C_i) \times R(C_j) \longrightarrow R(C_{i+j})$$

by $$(\alpha_i, \alpha_j) \longmapsto \alpha_i \cdot \alpha_j = (\alpha_i \otimes \alpha_j)^{W(C_{i+j})},$$

i.e. the induced representation where the tensor product

$$\alpha_i \otimes \alpha_j \in R(C_i) \otimes_Z R(C_j)$$

is considered as an element of the representation ring of $W(C_i) \times W(C_j)$ in the usual way.

Lemma 5.1.1

If H and K are subgroups of the finite groups J and L respectively, and $\chi \in R(H)$, $\psi \in R(K)$ where R(G) denotes the representation ring of G, then

$$\chi^J \otimes \psi^L = (\chi \otimes \psi)^{J \times L} \text{ in } R(J \times L)$$

Proof Considering χ and ψ as generalised characters we have

$$(\chi^J \otimes \psi^L)(x,y) = \chi^J(x)\psi^L(y)$$

$$= \frac{1}{|H|}\left\{\sum_{\substack{j\in J \\ j^{-1}xj\in H}} \chi(j^{-1}xj)\right\}\frac{1}{|K|}\left\{\sum_{\substack{s\in L \\ s^{-1}ys\in K}} \psi(s^{-1}ys)\right\}$$

$$= \frac{1}{|H \times K|} \sum_{\substack{(j,s)\in J\times L \\ (j,s)^{-1}(x,y)(j,s)\in H\times K}} (\chi\otimes\psi)((j,s)^{-1}(x,y)(j,s))$$

$$= (\chi\otimes\psi)^{J\times L}(x,y).$$

Proposition 5.1.2

If we define $R(C) = \bigoplus_{i\geq 0} R(C_i)$, a direct sum of additive abelian groups, then the outer product makes $(R(C),.)$ a graded commutative ring with identity.

Note "commutative" means that $x.y = y.x$ for all x and $y \in R(C)$.

Proof By the linearity properties of tensor products and induced representations, only associativity and commutativity of the product require proof.

Associativity: Let χ, ψ, τ be in $R(C_i)$, $R(C_j)$ and $R(C_k)$ respectively.

$$(\chi.\psi).\tau = ((\chi\otimes\psi)^{W(C_{i+j})}\otimes\tau)^{W(C_{i+j+k})} \text{ by definition}$$

$$= (((\chi\otimes\psi)\otimes\tau)^{W(C_{i+j})\times W(C_k)})^{W(C_{i+j+k})} \text{ by Lemma 5.1.1}$$

with $H = W(C_i) \times W(C_j)$, $J = W(C_{i+j})$, $K = L = W(C_k)$,

$$= (\chi\otimes\psi\otimes\tau)^{W(C_{i+j+k})} \text{ by the transitivity of induction}$$

$$= \chi.(\psi.\tau) \text{ similarly.}$$

Commutativity: Let χ and ψ be as above. The permutation $1\mapsto i+1$, $2\mapsto i+2$, ..., $j\mapsto i+j$, $j+1\mapsto 1$, $j+2\mapsto 2$, ..., $i+j\mapsto i$ of the set $\{1,2,\ldots,i+j\}$ induces an inner automorphism of $W(C_{i+j})$ under which the two subgroups $W(C_i) \times W(C_j)$ and $W(C_j) \times W(C_i)$ are conjugate. This automorphism gives the equation

$$\chi.\psi = \psi.\chi \ .$$

Let S_n denote the symmetric group consisting of all permutations of the set $\{1,2,\ldots,n\}$. Write $S_0 = \{1\}$, which may be considered as the permutations of the empty set. We embed $S_n \times S_m$ in S_{n+m} by letting S_n act on $\{1,\ldots,n\}$ and S_m on $\{n+1,\ldots,n+m\}$ in the obvious way. Let $R(S_n)$ be the representation ring of S_n, and define the outer product: $R(S_n) \times R(S_m) \longrightarrow R(S_{n+m})$ by

$$(\alpha_n, \alpha_m) \longmapsto \alpha_n . \alpha_m = (\alpha_n \otimes \alpha_m)^{S_{n+m}},$$

in a similar fashion to that used for R(C).

Proposition 5.1.3

If we define $R(S) = \bigoplus_{i \geq 0} R(S_i)$, a direct sum of additive abelian groups, then the outer product makes $(R(S),.)$ a graded commutative ring with identity.

Proof Almost identical to that of Proposition 5.1.2.

Notes 1. The construction for R(S) is described in detail in Donald Knutson: "λ-rings and the representation theory of the symmetric group", Springer Lecture Notes 308. See also section 1 of M.F. Atiyah:"Power Operations in K-Theory", Quart. J. Math. Oxford (2), 17 (1966), pages 165-93.

2. Both the rings R(S) and R(C) are used implicitly (without being explicitly defined) in W. Specht's papers on the representations of the symmetric and hyperoctahedral groups. (These groups are just S_n and $W(C_n)$ in our notation). In fact the map

$$R(C) \longrightarrow \mathbb{Q}[s_1,s_2,\ldots,t_1,t_2,\ldots] \text{ given by}$$

$$\chi \longmapsto \text{characteristic of } \chi$$

in the notation of Specht's paper on the hyperoctahedral group is a ring monomorphism.

Next we examine the relationships between the structures of the two graded rings R(S) and R(C). We may define two group homomorphisms $\beta : W(C_n) \longrightarrow S_n$ and $\gamma : S_n \longrightarrow W(C_n)$ for all n

by $(\beta t)(i) = |t(i)|$ and $(\gamma u)(j) = \mathrm{sgn}(j)u(|j|)$.
It is easily verified that this gives $W(C_n)$ the structure of a semi-direct product, i.e. the following sequence of groups is exact

$$1 \to (Z_2)^n \hookrightarrow W(C_n) \underset{\gamma}{\overset{\beta}{\rightleftarrows}} S_n \to 1$$

with $\beta\gamma$ = the identity map on S_n, where $(Z_2)^n = \mathrm{Ker}(\beta)$ is the elementary abelian 2-group on the generators $(1,-1),(2,-2),\ldots,(n,-n)$ and (a,b) denotes the transposition interchanging a and b but fixing everything else.

Then β induces a homomorphism of graded abelian groups $\beta^*: R(S) \longrightarrow R(C)$ by $\beta^*(\chi) = \chi\beta$, (composition of maps where χ is considered, for example, as the character map from S_n to the complex numbers).

We may define a one-dimensional representation $\xi_n: W(C_n) \longrightarrow C^*$, where C^* is the multiplicative group of non-zero complex numbers, by $\xi_n(t) = \frac{1}{n!}\prod_{i=1}^{n} t(i) = (-1)^{p(t)}$ where $p(t)$ is the number of i with $i > 0 > t(i)$. This induces another graded abelian group homomorphism $\xi_*: R(C) \longrightarrow R(C)$ where $\xi_*(\chi) = \xi_n\chi$, the ordinary internal product in the representation ring $R(C_n)$, (i.e. pointwise multiplication of characters as functions from the group to the complex numbers, if we consider $R(C_n)$ as the character ring of $W(C_n)$).

Lemma 5.1.4

$\beta^*: R(S) \longrightarrow R(C)$ and $\xi_*: R(C) \longrightarrow R(C)$ are both graded ring homomorphisms.

Proof Straightforward, after noting that $\beta_{i+j}|W(C_i) \times W(C_j) = \beta_i \times \beta_j$ and $\xi_{i+j}|W(C_i) \times W(C_j) = \xi_i \otimes \xi_j$.

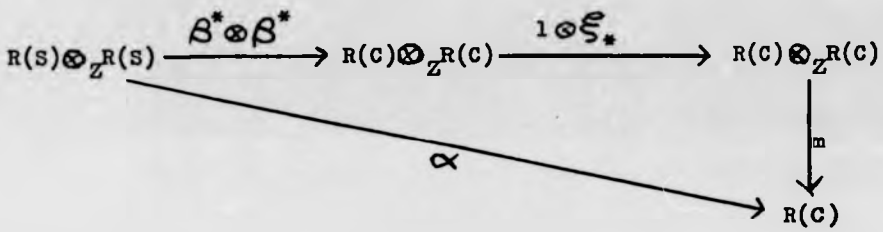
Write $(\ ,\ )$ for the usual scalar product of characters in $R(S_i)$, (i.e. $(\chi,\psi) = \frac{1}{|S_i|}\sum_{x \in S_i} \overline{\chi(x)}\psi(x)$, so that the absolutely

irreducible characters form an orthonormal basis). Extend this to a Z-bilinear positive definite scalar product on R(S) by defining

$$(\sum_{i=0}^{\infty}\chi_i, \sum_{j=0}^{\infty}\psi_j) = \sum_{k=0}^{\infty}(\chi_k,\psi_k) \text{ for all } \chi_i,\psi_i \in R(S_i),$$

where (of course) all the sums are finite. Then $R(S)\otimes_Z R(S)$ has the (usual) graded ring structure with k^{th} grading $\bigoplus_{i+j=k} R(S_i)\otimes_Z R(S_j)$, and product induced by $(x\otimes y).(u\otimes v) = (x.u)\otimes(y.v)$. Give it the (unique) scalar product induced by $(x\otimes y, u\otimes v) = (x,u)(y,v)$. Make all the analogous constructions for the ring R(C).

Define a map α to make the following diagram commute:



where m is the multiplication map:

$$m(\sum_{i,j} x_i\otimes y_j) = \sum_{i,j} x_i.y_j \quad .$$

Theorem 5.1.5

(i) α is a graded ring homomorphism: $R(S)\otimes_Z R(S) \longrightarrow R(C)$.

(ii) $(\alpha x,\alpha y) = (x,y)$ for all $x,y\in R(S)\otimes_Z R(S)$. (i.e. α is an isometry of inner product spaces over Z).

Proof (i) It is easily checked, with the aid of Lemma 5.1.4, that each of the horizontal and vertical maps in the diagram is a graded ring homomorphism. (In fact "$\otimes_Z$" is the coproduct in the category of graded commutative rings with identity and graded ring homomorphisms). Hence the composition, α, of these maps is also a graded ring homomorphism.

(ii) Since α is Z-linear, it is enough to prove (ii) in the case where $x = \chi\otimes\psi\in R(S_i)\otimes_Z R(S_j)$ and $y = \tau\otimes\varphi\in R(S_k)\otimes_Z R(S_m)$.

Case I: If $i+j \neq k+m$, then $(x,y) = 0 = (\alpha x,\alpha y)$ by definition.

Case II: Suppose $i+j = k+m$ but $j \neq m$ so that $(x,y) = 0$. By definition

$$\alpha_x = (\beta^*\chi \otimes \xi_j\beta^*\psi)^{W(C_{i+j})}, \quad \alpha_y = (\beta^*\tau \otimes \xi_m\beta^*\varphi)^{W(C_{i+j})},$$

and it is easy to see that the restriction of the character α_x to $(Z_2)^{i+j} = \mathrm{Ker}(\beta) \subseteq W(C_{i+j})$ contains only (linear) irreducible characters which take the value -1 on exactly j of the generators $(1,-1),\ldots,(i+j,-i-j)$ of $(Z_2)^{i+j}$. Hence the restrictions of α_x and α_y to this subgroup are disjoint, and so certainly $(\alpha_x,\alpha_y) = 0$.

Case III: $i = k$, $j = m$. By Z-linearity it is enough to consider the case where χ, ψ, τ and φ are all irreducible characters. However this (and indeed case II above) is a special case of Proposition 25a,b , pages 78–9 of J.P. Serre: "Representations lineaires des groupes finis", Hermann (2^{nd} Edition 1971).

Corollary 5.1.6

(i) If $\chi \in R(S_i)$ and $\psi \in R(S_j)$ are irreducible representations, then $\alpha(\chi\otimes\psi) \in R(C_{i+j})$ is also an irreducible representation.

(ii) Distinct sets (i,j,χ,ψ) give rise to distinct irreducible representations in $R(C)$. Every irreducible character in $R(C)$ arises in this way.

(iii) $\alpha : R(S)\otimes_Z R(S) \longrightarrow R(C)$ is a graded ring isomorphism.

Proof $(\alpha(\chi\otimes\psi),\alpha(\chi\otimes\psi)) = (\chi\otimes\psi,\chi\otimes\psi) = (\chi,\chi)(\psi,\psi) = 1$, and so $\alpha(\chi\otimes\psi) = \pm$ an irreducible representation, but its dimension is positive.

(ii) Distinctness follows from Theorem 5.1.5(ii). Let $\chi_1^{(i)}, \chi_2^{(i)},\ldots$ be a complete set of irreducible representations in $R(S_i)$ for each i, and write $\chi_{s,t}^{(i,j)} = \alpha(\chi_s^{(i)}\otimes\chi_t^{(j)})$, an irreducible representation in $R(C_{i+j})$. Then for each fixed n we have:

$$\sum_{i+j=n}\sum_{s,t} (\dim\chi_{s,t}^{(i,j)})^2 = \sum_{i+j=n}\sum_{s,t}\left\{[W(C_n):W(C_i)\times W(C_j)]\dim\chi_s^{(i)}\dim\chi_t^{(j)}\right\}^2$$

$$= \sum_{i+j=n} \left\{ \frac{(i+j)!^2}{(i!j!)^2} \sum_s (\dim \chi_s^{(i)})^2 \sum_t (\dim \chi_t^{(j)})^2 \right\}$$

$$= \sum_{i+j=n} \frac{n!^2}{(i!j!)^2} |S_i| |S_j| = n! \sum_{i+j=n} \frac{n!}{i!j!} = 2^n n!$$

$$= |W(C_n)| .$$

Hence this is a complete set of irreducible representations of $W(C_n)$. (iii) α is injective by Theorem 5.1.5(ii), surjective by (ii) above, and a graded ring homomorphism by Theorem 5.1.5(i).

Note Parts (i) and (ii) of the corollary are essentially the original results obtained independently by Specht and Young which enabled them to express the representation theory of $W(C_n)$ in terms of the well-known theory of the symmetric group. The introduction of the rings R(S) and R(C) makes it possible to state these results as a theorem about graded commutative rings with scalar products. In the next section we see that the ring structure of R(S) is completely determined by a classical result of Frobenius, and we use this to find the structure of R(C).

5.2 Partitions, ring structures, irreducible representations and Weyl subgroups

For $n \geqslant 0$ write $\lambda \vdash n$ if $\lambda = (0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n)$ is a sequence of non-negative integers with $\lambda_1+\lambda_2+\ldots+\lambda_n = n$. Given $\lambda \vdash n$, define $\bar{\lambda}_i = \lambda_i+i-1$ for $1 \leqslant i \leqslant n$, so that $0 < \bar{\lambda}_1 < \bar{\lambda}_2 < \ldots < \bar{\lambda}_n$, and define a polynomial in n variables with integer coefficients:

$$Q_\lambda(x_1,\ldots,x_n) = \sum_t \operatorname{sgn}(t) x_{\bar{\lambda}_1-t_1} x_{\bar{\lambda}_2-t_2} \ldots x_{\bar{\lambda}_n-t_n}$$

where $t = (t_1,\ldots,t_n)$ runs through all the permutations of the set $\{0,1,\ldots,n-1\}$ and we have the conventions: $x_0 = 1$ and $x_i = 0$ for $i < 0$. It is easy to show that this polynomial involves no non-zero terms in any x_i for $i > n$. (The polynomials Q_λ arise in the use of Schur functions in the representation theory of the symmetric groups, some of which is quoted below). The polynomial may also be written as a determinant:

$$Q_\lambda(x_1,\ldots,x_n) = \begin{vmatrix} x_{\lambda_1} & x_{\lambda_1-1} & x_{\lambda_1-2} & \cdots & x_{\lambda_1-n+1} \\ x_{\lambda_2+1} & x_{\lambda_2} & x_{\lambda_2-1} & \cdots & x_{\lambda_2-n+2} \\ x_{\lambda_3+2} & x_{\lambda_3+1} & x_{\lambda_3} & \cdots & x_{\lambda_3-n+3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{\lambda_n+n-1} & x_{\lambda_n+n-2} & x_{\lambda_n+n-3} & \cdots & x_{\lambda_n} \end{vmatrix}$$

$$= \det(a_{ij}) \text{ where } a_{ij} = x_{\lambda_i+i-j} \text{ for } 1 \leqslant i,j \leqslant n.$$

Theorem 5.2.1

(i) $R(S) = Z[X_1,X_2,\ldots]$, a polynomial ring in countably infinitely many (algebraically independent) variables, where $X_i = 1_{S_i} \in R(S_i)$ is the trivial one-dimensional representation of S_i, and $Z = R(S_0)$.

(ii) $X_iX_jX_k\ldots = (1_{S_i} \times S_j \times S_k \times \ldots)^{S_{i+j+k+\ldots}} \in R(S_{i+j+k+\ldots})$ with the obvius notation.

(iii) $R(C) = Z[Y_1,Y_2,\ldots,\xi_1,\xi_2,\ldots]$, another polynomial ring where

$Y_i = \alpha(X_i \otimes 1) = 1_{W(C_i)} \in R(C_i)$ and $\xi_i = \alpha(1 \otimes X_i) \in R(C_i)$ is the one-dimensional representation of $W(C_i)$ defined before Lemma 5.1.4.

(iv) $Y_i Y_j \ldots \xi_r \xi_s \ldots = (1_{W(C_i)} \times W(C_j) \times \ldots \otimes \xi_r \otimes \xi_s \otimes \ldots)^{W(C_{i+j..+r..})}$

(v) The irreducible representations in $R(S_n)$ are given by $\chi^{(\lambda)} = Q_\lambda(X_1, \ldots, X_n)$ for $\lambda \vdash n$. Distinct partitions λ give distinct representations.

(vi) The irreducible representations in $R(C_n)$ are given by $\chi^{(\lambda;\mu)} = \alpha(\chi^{(\lambda)} \otimes \chi^{(\mu)}) = Q_\lambda(Y_1, \ldots, Y_i) Q_\mu(\xi_1, \ldots, \xi_j)$ where $\lambda \vdash i$, $\mu \vdash j$ with $i+j = n$. Distinct pairs $(\lambda;\mu)$ give distinct representations.

Proof (ii) and (iv) follow by repeated applications of Lemma 5.1.1, noting that $1_H \otimes 1_K = 1_{H \times K}$.

(v) This is a theorem of Frobenius, "Ueber die Charaktere der symmetrischen Gruppe", SitzBer. Preuss. Akad., Berlin (1900), 516.

(i) $R(S_n)$ is (additively) the free abelian group generated by the irreducible representations of S_n. By (v) this group is also generated by the monomials $X_{\lambda_1} X_{\lambda_2} \ldots X_{\lambda_n}$ for $\lambda \vdash n$. (Inspection of the definition of $Q_\lambda(X_1, \ldots, X_n)$ reveals that all the monomials occurring in it also belong to $R(S_n)$). The number of such monomials is equal to the number of partitions of n, which is the same as the number of conjugacy classes in S_n, and hence the same as the number of inequivalent irreducible representations. Hence the monomials must freely generate the additive abelian group $R(S_n)$. This shows that $X_1, \ldots$ generate R(S) as a Z-algebra, and that they are algebraically independent over Z.

(For a fairly short direct proof of (i), see corollary 1.5 of M.F. Atiyah: "Power operations in K-theory", Quart. J. Math. Oxford (2), 17 (1966), pages 165-93).

(iii) Immediate from Corollary 5.1.6(iii) and the definition of α.

(vi) follows from (v), Corollary 5.1.6(i),(ii) and the fact that α

$Y_i = \alpha(X_i \otimes 1) = 1_{W(C_i)} \in R(C_i)$ and $\xi_i = \alpha(1 \otimes X_i) \in R(C_i)$ is the one-dimensional representation of $W(C_i)$ defined before Lemma 5.1.4.

(iv) $Y_i Y_j \ldots \xi_r \xi_s \ldots = (1_{W(C_i)} \times W(C_j) \times \ldots \otimes \xi_r \otimes \xi_s \otimes \ldots)^{W(C_{i+j..+r..})}$

(v) The irreducible representations in $R(S_n)$ are given by $\chi^{(\lambda)} = Q_\lambda(X_1, \ldots, X_n)$ for $\lambda \vdash n$. Distinct partitions λ give distinct representations.

(vi) The irreducible representations in $R(C_n)$ are given by $\chi^{(\lambda;\mu)} = \alpha(\chi^{(\lambda)} \otimes \chi^{(\mu)}) = Q_\lambda(Y_1, \ldots, Y_i) Q_\mu(\xi_1, \ldots, \xi_j)$ where $\lambda \vdash i$, $\mu \vdash j$ with $i+j = n$. Distinct pairs $(\lambda;\mu)$ give distinct representations.

Proof (ii) and (iv) follow by repeated applications of Lemma 5.1.1, noting that $1_H \otimes 1_K = 1_{H \times K}$.

(v) This is a theorem of Frobenius, "Ueber die Charaktere der symmetrischen Gruppe", SitzBer. Preuss. Akad., Berlin (1900), 516.

(i) $R(S_n)$ is (additively) the free abelian group generated by the irreducible representations of S_n. By (v) this group is also generated by the monomials $X_{\lambda_1} X_{\lambda_2} \ldots X_{\lambda_n}$ for $\lambda \vdash n$. (Inspection of the definition of $Q_\lambda(X_1, \ldots, X_n)$ reveals that all the monomials occurring in it also belong to $R(S_n)$). The number of such monomials is equal to the number of partitions of n, which is the same as the number of conjugacy classes in S_n, and hence the same as the number of inequivalent irreducible representations. Hence the monomials must freely generate the additive abelian group $R(S_n)$. This shows that $X_1, \ldots$ generate R(S) as a Z-algebra, and that they are algebraically independent over Z.

(For a fairly short direct proof of (i), see corollary 1.5 of M.F. Atiyah: "Power operations in K-theory", Quart. J. Math. Oxford (2), 17 (1966), pages 165-93).

(iii) Immediate from Corollary 5.1.6(iii) and the definition of α.

(vi) follows from (v), Corollary 5.1.6(i),(ii) and the fact that α

is a ring homomorphism.

Given two partitions $\alpha \vdash r$ and $\beta \vdash s$, define the Weyl subgroup

$$W(\alpha;\beta) = W(C_{\alpha_1}) \times \ldots \times W(C_{\alpha_r}) \times S_{\beta_1} \times \ldots \times S_{\beta_s} \subseteq W(C_{r+s})$$

by letting $W(C_{\alpha_1})$ act on $\{\pm 1,\ldots,\pm\alpha_1\}$, $W(C_{\alpha_2})$ act on $\{\pm(\alpha_1+1),\ldots,\pm(\alpha_1+\alpha_2)\}$, ... , and then letting S_{β_1} act on $\{\pm(r+1),\ldots,\pm(r+\beta_1)\}$ by permuting the set $\{r+1,\ldots,r+\beta_1\}$ in the obvious way and preserving the "+" or "−" signs, and letting the later symmetric groups S_{β_i} act on the obvious sets in a similar fashion. In particular, if $r = 0$ and $\beta = (0,\ldots,0,n) \vdash n$, we get the Weyl subgroup $W(0;n) = S_n \subseteq W(C_n)$. (Equivalently we may consider the symmetric groups as embedded in $W(C_n)$ via the homomorphism γ described before Lemma 5.1.4).

Lemma 5.2.2

$$(1_{S_n})^{W(C_n)} = \chi^{(n;0)} + \chi^{(n-1;1)} + \ldots + \chi^{(n-i;i)} + \ldots + \chi^{(0;n)}$$

for all $n \geqslant 1$, where "i" denotes the partition $\lambda = (0,\ldots,0,i) \vdash i$.

Proof The n+1 characters on the right are distinct and irreducible. Also $\chi^{(n-i;i)} = Q_{(n-i)}(X_1,\ldots,X_{n-i})Q_{(i)}(\xi_1,\ldots,\xi_i)$ in $R(C)$

$$= X_{n-i}\xi_i \text{ by inspection}$$

$$= (1_{W(C_{n-i})} \otimes \xi_i)^{W(C_n)}.$$

Thus

$$\sum_{i=0}^{n} \dim \chi^{(n-i;i)} = \sum_{i=0}^{n} \left[W(C_n):W(C_{n-i}) \times W(C_i)\right]$$

$$= \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} = 2^n = \left[W(C_n):S_n\right].$$

Hence it is enough to show that each of the irreducible representations on the right of the equation is a constituent of the left-hand side.

Since $(Z_2)^n \subseteq W(C_{n-i}) \times W(C_i)$ and $W(C_M) = (Z_2)^n S_n$, the "Mackey decomposition" gives

$$\left((1_{S_n})^{W(C_n)}, \chi^{(n-i;i)} \right)_{W(C_n)} = \left((1_{S_n})^{W(C_n)}, (1_{W(C_{n-i})} \otimes \xi_i)^{W(C_M)} \right)_{W(C_n)}$$

$$= ((1_{S_n})|H, (1_{W(C_{n-i})} \otimes \xi_i)|H)_H$$

$$= (1_H, 1_H)_H = 1 > 0, \text{ as required,}$$

where $H = S_n \cap (W(C_{n-i}) \times W(C_i)) = S_{n-i} \times S_i$, and $(\ ,\ )_K$ denotes the usual inner product for representations of the group K.

Theorem 5.2.3

(i) $R(C) = Z[Y_1, Y_2, \ldots, Z_1, Z_2, \ldots]$, a polynomial ring, where

$$Y_i = 1_{W(C_i)} \text{ and } Z_i = (1_{S_i})^{W(C_i)} \in R(C_i).$$

(ii) For all $n \geq 1$,

$$Z_n = Y_n + Y_{n-1}\xi_1 + \cdots + Y_{n-i}\xi_i + \cdots + Y_1\xi_{n-1} + \xi_n.$$

(iii) (S.J. Mayer: Ph.D. Thesis, Warwick 1971).

The set $(1_{W(\alpha;\beta)})^{W(C_n)}$: $W(\alpha;\beta)$ a Weyl subgroup of $W(C_n)$ is an integral basis of $R(C_n)$ = the character ring of $W(C_n)$

Proof (ii) is just a restatement of the preceding lemma. Using this we may solve recursively for ξ_n in terms of the Y_i and Z_j:

$$\xi_1 = Z_1 - Y_1, \text{ and } \xi_n = Z_n - Y_n - Y_{n-1}\xi_1 - \cdots - Y_1\xi_{n-1} \text{ for } n > 1.$$

So the Y_i and Z_j generate R(C) as a Z-algebra since the Y_i and ξ_j do. In particular the monomials $Y_{\alpha_1} \cdots Y_{\alpha_r} Z_{\beta_1} \cdots Z_{\beta_s} = (1_{W(\alpha;\beta)})^{W(C_n)}$ for $\alpha \vdash r$, $\beta \vdash s$, $r+s = n$, generate $R(C_n)$ as an abelian group. However this set has the same cardinality as the set of irreducible characters of $W(C_n)$. This proves (iii).

The proof of (i) is now similar to that of Theorem 5.2.1(i).

In the next section we solve the recurrence relation of

Theorem 5.2.3(ii) for ξ_n in terms of the Y_i and Z_j, and deduce an explicit formula for the irreducible characters in terms of the new integral basis of Theorem 5.2.3(iii).

5.3 Weyl and parabolic subgroups of $W(C_n)$

We keep the notations of sections 5.1 and 5.2. First we rewrite the irreducible representations $\chi^{(\lambda;\mu)}$ in terms of the new generators of $R(C)$ introduced in Theorem 5.2.3.

Lemma 5.3.1

$\xi_j = P_0Z_j + P_1Z_{j-1} + \cdots + P_{j-1}Z_1 + P_j$ for each $j \geq 0$, where $P_i \in R(C_i)$ is defined recursively by $P_0 = 1$ and $P_i = -\sum_{k=0}^{i-1} P_kY_{i-k}$ for $i > 0$. In particular, P_i is a polynomial with integer coefficients in $Y_1, Y_2, \ldots, Y_i$ only.

Proof We use induction on j. It is true for $j = 0, 1$. (By the convention introduced in 5.2, $\xi_0 = 1$). Suppose the result already proved for $\xi_0, \ldots, \xi_{j-1}$. Theorem 5.2.3(ii) gives

$$\xi_j = Z_j - \sum_{i=0}^{j-1} \xi_iY_{j-i} = P_0Z_j - \sum_{i=0}^{j-1} (P_0Z_i + P_1Z_{i-1} + \cdots + P_i)Y_{j-i} \quad \text{by the induction hypothesis,}$$

$$= P_0Z_j - \sum_{k=0}^{j-1} Z_k \Big(\sum_{i=k}^{j-1} P_{i-k}Y_{j-i} \Big) \text{ by reordering the summation}$$

$$= P_0Z_j - \sum_{k=0}^{j-1} Z_k \Big(\sum_{t=0}^{j-1-k} P_tY_{j-k-t} \Big) \text{ substituting } t = i-k$$

$$= P_0Z_j + \sum_{k=0}^{j-1} Z_kP_{j-k} \text{ by the recurrence relation for the P's.}$$

We now solve this recurrence relation for P_i, and discover that the solution involves some of the polynomials Q_λ defined in 5.2.

Lemma 5.3.2

$$P_i = (-1)^i Q_{(1^i)}(Y_1, Y_2, \ldots, Y_i)$$

where $(1^i) \vdash i$ is the partition $(1, \ldots, 1)$ of i.

Proof By the last lemma $P_0,\ldots,P_i$ are the solutions of the following simultaneous linear equations in R(C):

$$\begin{aligned}
&P_0 && && && = 1\\
&Y_1P_0 + && P_1 && && = 0\\
&Y_2P_0 + && Y_1P_1 + && P_2 && = 0\\
&\vdots && \vdots && \vdots && \vdots\\
&Y_iP_0 + && Y_{i-1}P_1 + && Y_{i-2}P_2 + \cdots + Y_1P_{i-1} + P_i && = 0
\end{aligned}$$

Since R(C) is a commutative ring we may apply Cramer's rule to these:

$$P_i \begin{vmatrix} 1 & 0 & 0 & \cdots & 0\\ Y_1 & 1 & 0 & \cdots & 0\\ Y_2 & Y_1 & 1 & & 0\\ \vdots & \vdots & & & \vdots\\ \vdots & \vdots & & 1 & 0\\ Y_i & Y_{i-1} & \cdots & Y_1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & \cdots & 0 & 1\\ Y_1 & 1 & \cdots & 0 & 0\\ Y_2 & \vdots & & \vdots & 0\\ \vdots & \vdots & & \vdots & \vdots\\ \vdots & \vdots & & 1 & 0\\ Y_i & Y_{i-1} & \cdots & Y_1 & 0 \end{vmatrix}$$

The left-hand determinant is just 1, and expanding by the elements of the last column shows that the second determinant is

$$(-1)^i \begin{vmatrix} Y_1 & 1 & 0 & \cdots & 0\\ Y_2 & Y_1 & 1 & & 0\\ \vdots & \vdots & & & \vdots\\ \vdots & \vdots & & & 1\\ Y_i & Y_{i-1} & \cdots & & Y_1 \end{vmatrix} = (-1)^i Q_{(1^i)}(Y_1,\ldots,Y_i) \text{ as required.}$$

Summing up, we have:

Theorem 5.3.3

The irreducible representations in $R(C_n)$ are the polynomials $\chi^{(\lambda;\mu)} = Q_\lambda(Y_1,\ldots,Y_i)Q_\mu(\xi_1,\ldots,\xi_j)$ for $\lambda \vdash i$, $\mu \vdash j$, $i+j = n$, where $\xi_j = P_0Z_j + P_1Z_{j-1} + \cdots + P_{j-1}Z_1 + P_j$, and $P_k = (-1)^k Q_{(1^k)}(Y_1,\ldots,Y_k)$, with (1^k) being the partition $(1,\ldots,1)$ of k. The monomial $Y_{\lambda_1}\cdots Y_{\lambda_r}Z_{\mu_1}\cdots Z_{\mu_s}$ is just $(1_{W(\lambda;\mu)})^{W(C_n)}$.

Proof This is just a restatement of earlier results.

This gives a complete solution to the problem of expressing the irreducible characters of $W(C_n)$ as Z-linear combinations of the induced principal characters $\{(1_H)^{W(C_n)} : H \text{ a Weyl subgroup of } W(C_n)\}$. However, for many problems involving the relationship between a Weyl group and various associated Chevalley groups, the most interesting class of subgroups of $W(C_n)$ is a proper subset of the Weyl subgroups.

If $\mu \vdash s \leq n$, then the parabolic subgroup $W(\mu)$ of $W(C_n)$ is is defined to be the Weyl subgroup $W(n-s;\mu)$, i.e.

$$W(\mu) = W(C_{n-s}) \times S_{\mu_1} \times S_{\mu_2} \times \ldots \times S_{\mu_s} \subseteq W(C_n).$$

Then $(1_{W(\mu)})^{W(C_n)} = Y_{n-s} Z_{\mu_1} \ldots Z_{\mu_s}$. We call a monomial $Y_{\alpha_1} \ldots Y_{\alpha_r} Z_{\beta_1} \ldots Z_{\beta_s}$ in $R(C)$ "parabolic" if at most one $\alpha_i > 0$, (i.e. if its total degree in the polynomial arguments $Y_1, Y_2, \ldots$ is 0 or 1, or equivalently if $W(\alpha;\beta)$ is a parabolic subgroup. We are assuming, of course, that the monomial is non-zero).

Note The terms "parabolic subgroup" and "Weyl subgroup" originate in the theory of Weyl groups, but the representation theory that we are developing here does not require a detailed understanding of most of their properties. Clearly the theory above would be unaltered if the Weyl subgroups were replaced by conjugate subgroups. For our purposes a parabolic subgroup may be regarded as one generated by a subset of the transpositions $(1,-1)$, $(1,2)$, $(2,3)$, $\ldots$, $(n-1,n)$ or any subgroup conjugate to one of these. (The transposition (i,j) in S_n is regarded as the element of $W(C_n)$ which interchanges i and j, and also interchanges $-i$ and $-j$). A Weyl subgroup may be regarded as one generated by any subset of the transpositions $(1,-1),(2,-2),\ldots,(n,-n),(1,2),(2,3),\ldots,(n-1,n)$, or as any subgroup conjugate to one of these.

We wish to determine which irreducible representations of $W(C_n)$ may be written as Z-linear combinations of parabolic monomials.

Interest in such representations is motivated by results of Steinberg, Curtis and Solomon which we describe later. In order to handle this problem we need some technical results on the polynomials Q_λ and algebraic independence over the integers.

Lemma 5.3.4

Let $x_1, x_2, \ldots$ be algebraically independent over the integers Z. Let $\lambda \vdash r$ and $\mu \vdash s$. Then

(i) The monomial $x_{\lambda_1} x_{\lambda_2} \ldots x_{\lambda_r}$ has coefficient +1 in the expansion of the polynomial $Q_\lambda(x_1, \ldots, x_r)$.

(ii) If $\lambda \neq \mu$, then $Q_\lambda(x_1, \ldots, x_r)$ and $Q_\mu(x_1, \ldots, x_s)$ are linearly independent over Z.

Proof (i) Recall the definition of Q_λ at the beginning of 5.2. If $x_{\lambda_1} \ldots x_{\lambda_r} = x_{\bar{\lambda}_1 - t_1} \ldots x_{\bar{\lambda}_r - t_r}$, where $\bar{\lambda}_i = \lambda_i + i - 1$ and $(t_1, \ldots, t_r)$ is a permutation of $(0, 1, \ldots, r-1)$, we must have $t_1 = 0$, for otherwise $\bar{\lambda}_1 - t_1 < \lambda_1$ which is the least of the λ_i, contradicting the equality of the two monomials. Similarly $t_2 = 1$, etc.

(ii) This is necessary by Theorem 5.2.1(v) since the irreducible characters of a finite group are linearly independent over Z. (Obviously a more elementary proof of this is possible, but it would be much longer).

Lemma 5.3.5

(i) In R(C) the term not involving $Z_1, Z_2, \ldots$ in the expansion of the polynomial $Q_\lambda(\xi_1, \ldots, \xi_r)$ in powers of the Y_i and Z_j for $\lambda \vdash r$ is $Q_\lambda(P_1, \ldots, P_r) \in R(C_r)$.

(ii) The elements $P_1, P_2, \ldots$ of R(C) are algebraically independent over Z.

Proof (i) Immediate upon setting $Z_j = 0$ for $j \geqslant 1$ and using Lemma 5.3.1.

(ii) $Y_1, Y_2, \ldots$ are algebraically independent over Z, and

$P_1 = -Y_1$, $P_i = -(P_0Y_i+P_1Y_{i-1}+\ldots+P_{i-1}Y_1)$ for $i \geq 1$. Hence

$$Z[P_1,\ldots,P_r] = Z[Y_1,\ldots,Y_r] \text{ for all } r \geq 1.$$

The result now follows by the standard theory of transcendence degree.

Theorem 5.3.6

For each $n > 0$ there are at most two irreducible representations of $W(C_n)$ which may be written as Z-linear combinations of parabolic monomials in $R(C)$.

Proof Suppose $\chi^{(\lambda;\mu)} = Q_\lambda(Y_1,\ldots,Y_i)Q_\mu(\xi_1,\ldots,\xi_j)$ is a linear combination of parabolic monomials. Since its total degree in the Y_j is at most one, at most one $\lambda_t \neq 0$ by Lemma 5.3.4(i). Hence $\lambda = (0,\ldots,0,i) \vdash i$ and $Q_\lambda(Y_1,\ldots,Y_i) = Y_i$.

Similarly, by Lemma 5.3.5(i), $Q_\mu(P_1,\ldots,P_j) \in R(C_j)$ has total degree at most one in $Y_1,\ldots,Y_j$. Hence $Q_\mu(P_1,\ldots,P_j) = mY_j$ for some integer m. For each j this is possible for at most one $\mu \vdash j$ by Lemma 5.3.5(ii) and Lemma 5.3.4(ii).

Thus the term independent of $Z_1,Z_2,\ldots$ in $\chi^{(\lambda;\mu)}$ is mY_iY_j, which has total degree at most one in $Y_1,\ldots$ only if i or j = n. In each case there is at most one possibility for the pair $(\lambda;\mu)$.

Corollary 5.3.7

For $n > 0$, the only irreducible characters of $W(C_n)$ which may be written as Z-linear combinations of the characters of the form $(1_P)^{W(C_n)}$ for P a parabolic subgroup are $\chi^{(n;0)}$ and $\chi^{(0;1,\ldots,1)}$, both of which are one-dimensional.

Proof $\chi^{(1;0)}$ = the principal character by inspection. The other character is of the required form by a theorem of Solomon: "The orders of the finite Chevalley groups", J. Algebra 3 (1966), pages 376-393, Theorem 2.

Notes 1. One of the reasons for interest in this result is the following theorem of C.W. Curtis: "The Steinberg character of a finite group with a (B,N)-pair", J. Algebra 4 (1966),pages 433-441, Theorem 1:

Let G be a finite group with a (B,N)-pair and Weyl group W. Suppose χ is an irreducible character of W such that

$$\chi = \sum_{J \subseteq R} n_J (1_{W_J})^W, \ n_J \text{ integers.}$$

Then $\psi = \sum_{J \subseteq R} n_J (1_{G_J})^G$ is a generalised character of G such that $\pm \psi$ is an irreducible character.

(The detailed meaning of the symbols does not matter here. It is enough to know that the W_J are precisely the parabolic subgroups of W described above for $W = W(C_n)$, and that the G_J are the parabolic subgroups of G which correspond to the W_J under the "Bruhat decomposition").

Two characters always arise in this way. They are the principal character 1_G and the Steinberg character, which in our case corresponds to $\chi^{(0;1,\ldots,1)}$. We have shown that if $W = W(C_n)$ then only these two cases occur.

2. In the case of the Weyl groups $W(A_n)$ which are isomorphic to the symmetric groups, the concepts of Weyl subgroup and parabolic subgroup coincide, and in fact Frobenius' result, (Theorem 5.2.1(v) above), shows that every character of the Weyl group can be written as a Z-linear combination of characters induced from the principal characters of parabolic subgroups. Steinberg used this to find a large number of irreducible characters of the general linear groups over finite fields. See: "A geometric approach to the representations of the full linear group over a Galois field", Trans. Amer. Math. Soc. 71, 274.

Bibliography

1. Artin, E. : "Geometric algebra", Interscience (1957)
2. Atiyah, M.F. : "Power operations in K-theory", Quart. J. Math. Oxford (2), 17 (1966), 165-93
3. Birch, B.J. : "Cyclotomic fields and Kummer extensions" in [8]
4. Borel, A. : "Linear algebraic groups", Benjamin (1969)
5. Borel, A. et al. : "Seminar on algebraic groups and related finite groups", Springer-Verlag, Lecture Notes 131 (1970)
6. Carter, R.W. : "Simple groups of Lie type", Wiley (1972)
7. Carter, R.W. : "Conjugacy classes in the Weyl group" in [5]
8. Cassels, J.W.S & Frohlich, A. (ed) : "Algebraic number theory", Academic Press (1967)
9. Cikunov, I.K. : "The structure of isometric transformations of a symplectic or orthogonal vector space", Soviet Math. Dokl. 6 (1965), 1479-81
10. Curtis, C.W. : "The Steinberg character of a finite group with a (B,N)-pair", J. Algebra 4 (1966), 433
11. Curtis, C.W. : "Chevalley groups and related topics" in M.B. Powell & G. Higman (ed) : "Finite simple groups", Academic Press (1971)
12. Curtis, C.W. & Reiner, I. : "Representation theory of finite groups and associative algebras", Interscience (1962)
13. Dickson, L.E. : "Linear groups", Teubner (1901)
14. Dieudonné, J. : "La géométrie des groupes classiques", Springer-Verlag, 3rd edition (1971)
15. Feit, W. : "Characters of finite groups", Benjamin (1967)
16. Frobenius, G. : "Ueber die Charaktere der symmetrischen Gruppe", SitzBer. Preuss. Akad., Berlin (1900), 516
17. Green, J.A. : "The characters of the finite general linear groups", Trans. A.M.S. 80 (1955), 402-447

18. Greenberg, M.J. : "Lectures on forms in many variables", Benjamin (1969)

19. Huppert, B. : "Endliche Gruppen I", Springer-Verlag (1967)

20. Kerber, A. : "Representations of permutation groups I", Springer-Verlag, Lecture Notes 240 (1971)

21. Knutson, D. : "λ-rings and the representation theory of the symmetric group", Springer-Verlag, Lecture Notes 308 (1973)

22. Lang, S. : "Algebra", Addison-Wesley (1965)

23. Mayer, S.J. : Ph. D. Thesis, Warwick (1971)

24. Milnor, J. : "On isometries of inner product spaces", Inventiones Math. 8 (1969) 83-97

25. O'Meara, O.T. : "Introduction to quadratic forms", Springer-Verlag (1963)

26. Robinson, G. de B. : "Representation theory of the symmetric group", Edinburgh University Press (1961)

27. Rutherford, D.E. : "Substitutional Analysis", Edinburgh University Press (1948)

28. Serre, J.P. : "Representations lineaires des groupes finis", Hermann (1967)

29. Solomon, L. : "The orders of the finite Chevalley groups", J. Algebra 3 (1966) 376

30. Specht, W. : "Darstellungstheorie der Hyperoktaedergruppe", M.Z. 42 (1937) 629-40

31. Springer, T.A. : "Characters of special groups" in [5]

32. Springer, T.A. & Steinberg, R. : "Conjugacy classes" in [5]

33. Srinivasan, B. : "The characters of the finite symplectic group Sp(4,q)", Trans. A.M.S. 131 (1968), 488-525

34. Steinberg, R. : "A geometric approach to the representations of the full linear groupover a Galois field", Trans. A.M.S. 71 (1951) 274

35. Wall, G.E. : "On the conjugacy classes in the unitary, symplectic

and orthogonal groups", J. Austr. Math. Soc., 3 (1963) 1-62

36. Weyl, H. : "The classical groups", Princeton (1939)

37. Weyl, H. : "Algebraic theory of numbers", Princeton (1940)

38. Young, A. : "Quantative substitutional analysis", P.L.M.S. (nine papers between 1901 and 1952)

39. Zariski, O. & Samuel, P. : "Commutative Algebra", Van Nostrand (1958)

Key to Notations

Most of the notations in this thesis are explained as they are introduced. For the symbols most commonly needed we give below either a definition or the number of the section in which it is introduced.

Z = the ring of integers

Q = the field of rational numbers

$F[x]$ = the ring of polynomials in one variable x over the field F

R^* = the group of units of the (commutative) ring R

$\det(s)$ = the determinant of the linear map s

$(a_1,\ldots,a_n)$, $[a_1,\ldots,a_n]$, $a \equiv b \bmod(c)$, $\phi(m)$:	See 1.1
$\mathrm{Gal}(L/K)$, $K(\sqrt[m]{1})$, L^G:	See 1.2
$W(k)$, Q_p, Z_p, Q_q, Z_q, G_{ram}, G_{Fr}:	See 1.3
$\Gamma(G)$, $K(G)$, $\Gamma(G,K)$:	See 2.1
$GL(V)$, $O(V,f)$, $CO(V,f)$, $O(V)$, $CO(V)$, $Sp(V)$, $CSp(V)$, $U(V)$, $CU(V)$, $M(f)$, $\lambda(s)$:	See 3.1
$\mathrm{Auto}(G)$, $\mathrm{Aut}(V)$, VS, $V_s(p)$:	See 3.2
IPS, $V_s(\{p,p^*\})$:	See 3.3
$CIPS$, $\lambda(s)$:	See 3.4
$d(V)$:	See 4.2
$O_+(V)$, $O_-(V)$:	See 4.4
$SO(V)$:	See 4.6
$W(C_n)$, $R(C_n)$, $R(C)$, $R(S_i)$, $R(S)$:	See 5.1
$\lambda \vdash n$, $Q_\lambda(x_1,\ldots,x_n)$, $\chi^{(\lambda)}$, $\chi^{(\lambda;\mu)}$, $W(\alpha;\beta)$:	See 5.2